ARL-TR-7212 ● FEB 2015

US Army Research Laboratory

Investigating Surface Bias Errors in the Weather Research and Forecasting (WRF) Model using a Geographic Information System (GIS)

by Jeffrey A Smith, Theresa A Foley, John W Raby, and Brian Reen

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

ARL-TR-7212 ● FEB 2015

US Army Research Laboratory

Investigating Surface Bias Errors in the Weather Research and Forecasting (WRF) Model using a Geographic Information System (GIS)

by Jeffrey A Smith, Theresa A Foley, John W Raby, and Brian Reen
Computational and Information Sciences Directorate, ARL

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED (From - To)
Feb 2015	Final	Jan–Dec 2014

4. TITLE AND SUBTITLE

Investigating surface bias errors in the Weather Research and Forecasting (WRF) Model using a Geographic Information System (GIS)

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Jeffrey A Smith, Theresa A Foley, John W Raby, and Brian Reen

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

U.S. Army Research Laboratory
Computational and Information Sciences Directorate
Battlefield Environment Division (ATTN: RDRL-CIE-M)
White Sands Missile Range, NM 88002-5501

8. PERFORMING ORGANIZATION REPORT NUMBER

ARL-TR-7212

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The Weather Research and Forecasting Model (WRF) is a numerical weather prediction model maintained by the National Center for Atmospheric Research (NCAR). The US Army Research Laboratory (ARL) has developed the Weather Running Estimate–Nowcast (WRE–N) model based on the Weather Research and Forecasting Model–Advanced Research WRF (WRF–ARW) to predict small-scale phenomena of interest to Warfighters, such as mountain/valley breezes and land/sea breezes. NCAR has developed the Model Evaluation Tools (MET) to evaluate the accuracy of WRF forecasts using observations of meteorological variables. The traditional use of MET calculates model performance over the entire model domain. High-resolution modeling requires more focused verification at subdomain levels. A Geographical Information System (GIS) enables consideration of terrain variables in model assessments using a location-based approach. We discuss our GIS approach to verify WRF-ARW with a 1-kilometer horizontal resolution inner domain centered near San Diego, California. We selected 5 case study days from February and March of 2012. The literature indicated that elevation is strongly correlated with meteorological parameters, which became our focus. We found that elevation accounts for a significant portion of the variance in the model error. The study demonstrated that a GIS can analyze spatially distributed point forecast errors over subdomains, to reveal the dependence of forecast errors on elevation.

15. SUBJECT TERMS

WRF, model validation, model verification, weather forecasting, Geographic Information System, GIS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE	UU	88	Jeffrey A Smith
Unclassified	Unclassified	Unclassified			19b. TELEPHONE NUMBER (Include area code) (575) 678-1332

Contents

List of Figures

List of Tables

Acknowledgments

We offer our thanks to Mr Jeff Passner, Mr Bob Dumais, Dr Huaqing Cai, Dr Richard Penc, Lieutenant Colonel (LTC) John Olson, PhD, and Mr Jeff Johnson of the US Army Research Laboratory (ARL). Each of these individuals contributed to our work in varying degrees, and without their feedback we would not have progressed as far as we did.

The authors would also like to thank Olga Wilhelmi and Jennifer Boehnert, Geographic Information System (GIS) experts from the National Center for Atmospheric Research (NCAR), for their invaluable suggestions on how to import Model Evaluation Tools (MET) Point-Stat files into GIS. Without their advice, this project could not have been completed. The authors would also like to thank Jason Knievel, Barbara Brown, and Tressa Fowler of NCAR for their encouragement.

Many thanks go to Technical Editor Jenny Weathers and the Technical Publishing Branch at White Sands Missile Range (WSMR), New Mexico, for their consistently high standard of editing.

Executive Summary

The Weather Research and Forecasting Model (WRF) is a numerical weather prediction model that has been used for many applications, including My Weather Impacts Decision Aid (MyWIDA). WRF is maintained by the National Center for Atmospheric Research (NCAR), which has developed a suite of Model Evaluation Tools (MET) to evaluate the accuracy of WRF forecasts using observations of meteorological variables such as temperature, relative humidity, and wind. The traditional use of MET calculates model performance over the entire model domain.

For Soldiers in theater, the Army requires high-resolution weather forecasting to resolve atmospheric features with wavelengths on the order of 5 kilometers (km) or less. This requires models that operate on a model grid spacing of about 1 km or less in the finest, or most resolved, domain. To validate such high-resolution modeling against observations requires a more focused spatial and temporal approach over parts of the domain rather than the domain as a whole. With a Geographic Information System (GIS), researchers can now consider terrain type/slope, land use effects, and other spatial and temporal variables as explanatory metrics in model assessments. GIS techniques, when coupled with high-resolution point and gridded observation sets, allow location-based approaches that permit discovery of spatial and temporal scales where models do not sufficiently resolve the desired phenomena—for example, turbulence effects or mountain and lee waves.

In this technical report, we discuss our initial efforts in using GIS tools to study model errors in the Advanced Research WRF (WRF–ARW) with a 1-km horizontal grid spacing inner domain centered near San Diego, California. The San Diego area contains a mixture of urban, suburban, agricultural, and mountainous terrain types along with a rich array of observational data with which to illustrate our ability to conduct subdomain verification. We selected 5 case study days from February and March of 2012 with varied synoptic conditions.

A literature review indicated that elevation is strongly correlated with meteorological parameters, at least in a climatic sense; consequently we focused this report on elevation as the explanatory variable. We found that elevation accounts for a significant portion of the variance in the model error of surface temperature and relative humidity predictions. On average, elevation accounts for about 10% to 20% of the total variance. In some cases however, elevation explained more than 50% of the total variance, while in other cases elevation was not a significant explanatory variable at all. Overall we found that elevation

accounts for more variance during the afternoon hours and less variance during the early morning hours; however, we could not discern a clear diurnal pattern within the model errors.

In this study, we demonstrated the effectiveness of using a GIS to evaluate model performance at the subdomain level. A GIS can analyze spatially distributed point forecast errors and show the dependence of forecast errors on terrain characteristics at the spatial scales of phenomena of interest to Warfighters, such as mountain/valley breezes and land/sea breezes. The GIS techniques used in this study show considerable promise for demonstrating the relationships of high-resolution model errors to explanatory variables such as terrain as well as identifying subdomains where the spatial statistics are more homogeneous. Further analysis using other explanatory variables such as slope, land use, and distance from the coastline will be needed to draw firm conclusions about the model performance. We achieved our goal with this study, which was to develop a GIS analysis method that goes beyond the types of analysis and conclusions gained from domain-level, aggregate statistics.

1. Background

Weather has a significant impact on Army personnel, weapons, tactics, and operations; therefore, accurate weather forecasts can be a deciding factor in any conflict—large or small. A famous example is D-day in World War II when the Allied nations invaded Normandy, France. Only a few days each month had the timing of high and low tides and full moon conditions suitable for landing thousands of men on the beaches of Normandy. On June 4, 1944, high winds and heavy seas made a landing impossible; however, James Stagg of the Royal Air Force and his team of meteorologists predicted that the weather would improve on June 5, and General Eisenhower decided to go ahead with the invasion. The German meteorologists were not expecting the weather to break and were thus surprised by the invasion (Ross 2014).

As computing technology has advanced, the weather forecasting task, once the primary role of a human forecaster in theater, has shifted to computerized Numerical Weather Prediction (NWP) models. Scientists around the world have used the Weather Research and Forecasting (WRF) model extensively for many applications. In this study, we have used the Advanced Research version of WRF (Skamarock et al. 2008) that we abbreviate as WRF–ARW. WRF–ARW includes Four Dimensional Data Assimilation (FDDA) techniques that can be used to incorporate observations into the model so that forecast quality is improved (Deng et al. 2009; Stauffer and Seaman 1994). The US Army Research Laboratory (ARL) uses WRF–ARW as the core of its Weather Running Estimate–Nowcast (WRE–N) weather forecasting model.

The Army requires high-resolution weather forecasting to resolve atmospheric features with wavelengths on the order of 5 km or less, which imposes a requirement for NWP to operate on a model grid spacing on the order of 1 km or less in the finest, or most resolved, domain to resolve weather phenomena of interest to the Soldier in theater. The atmospheric flows of interest to the Army include mountain/valley breezes, sea breezes, and other flows induced by differences in land surface characteristics. High-resolution NWP forecasts need to be validated against observations before their outputs can be used by applications such as My Weather Impacts Decision Aid (MyWIDA) developed by Brandt et al. (2013). Weather forecast validation has always been of interest to the civilian and military weather forecasting community; see for example the reviews by Ebert et al. (2013) and Casati et al. (2008), or the guides by Jolliffe and Stephenson (2011) or Wilks (2011). The validation of the models, especially high-resolution NWP, has proven to be especially difficult when addressing small temporal and spatial

scales (National Research Council 2010) that characterize NWP for use in Army applications.

The WRF model is maintained by the National Center for Atmospheric Research (NCAR), which has also developed a suite of Model Evaluation Tools (MET) (NCAR 2013) to evaluate WRF–ARW performance. MET Point-Stat performs traditional grid-to-point verification, while MET Grid-Stat performs grid-to-grid neighborhood verification to account for the uncertainty inherent in high-resolution forecasting. The MET Method for Object-based Diagnostic Evaluation (MODE) has been used to develop techniques for object-based spatial verification of high-resolution forecast grids of continuous meteorological variables.

ARL has employed MET in prior assessments such as that of Raby et al. (2012) who evaluated 2 models to arrive at domain level conclusions about the various strengths and weaknesses of these models and their accuracies. While MET proved useful as an assessment tool for forecasts over a regional domain, they found its ability to illuminate model shortcomings on the spatial and temporal scales needed by battlefield commanders to be lacking. In another assessment Dumais et al. (2012) evaluated 3 models that produced 3-dimensional (3-D) wind fields to arrive at domain-level conclusions. Although Dumais et al. did not employ any of the MET in their analysis, their desire was to also identify model strengths and weaknesses, and again, they found their ability to draw spatial and temporal conclusions lacking.

In other related work, Johnson (2012) conducted a location-based analysis of WRF over an inland testing range and for specific model time bins for a single day of interest. In his analysis, Johnson categorized the terrain as belonging to Valleys, Plains, or Mountains and concluded that this categorization increased correlation coefficients for these categories. However, the differences when considering horizontal grid resolutions of 1 km versus 3 km were present only in the “Valley” and “Mountain” cases. Though differences in 1 km and 3 km were present (favoring 1 km improved accuracy), they were minimal. Johnson et al. (2014) extends and refines this result.

Our assertion (Smith et al. 2014a; Smith et al. 2014b; 2015; Smith et al. 2014c; Smith et al. 2013) is that high-resolution models require verification on temporal and spatial scales appropriate for the phenomena that are being forecast. Our principal thesis is that a Geographic Information System (GIS) has tools that allow us to divide the domain into more homogenous subdomains according to explanatory variables such as terrain and land use. Our ultimate goal is to make more accurate inferences about model performance as well as providing a better understanding of model strengths and weaknesses.

The purpose of this study is to develop GIS tools that can better evaluate model performance. Our study domain is located in the San Diego, California region, which we chose because of its varied terrain and large number of weather observing stations. We chose 5 case study days with varied synoptic conditions from February and March of 2012. We used the MET Point-Stat to calculate the model forecast value at each of the observing stations and analyzed the matched forecast–observation pairs using GIS tools.

2. Domain and Model

The ARL WRE–N (Dumais et al. 2013; Dumais et al. 2004) has been designed as a convection-allowing application of the WRF–ARW model (Skamarock et al. 2008) with an observation-nudging FDDA option (Deng et al. 2009; Liu et al. 2005). For this investigation, we configured WRE–N in a multinest (i.e., 9/3/1 km) configuration to produce a fine inner mesh with 1-km grid spacing and leveraged an external global model for cold-start initial conditions and time-dependent lateral boundary conditions for the outermost nest. For ARL development and testing, this global model has been the National Center for Environmental Prediction's Global Forecast System (GFS) model (Environmental Modeling Center 2003). The WRE–N is envisioned to be a rapid-update cycling application of WRF–ARW with FDDA and optimally could refresh itself at intervals up to hourly (dependent upon the observation network) (Dumais et al. 2012; Dumais and Reen 2013).

For this particular experiment, the model was run with a base time of 1200 coordinated universal time (UTC) and generated output for each hour from 1200 UTC to 1200 UTC of the following day for a total of 25 model outputs on each of 5 days in February and March of 2012. The modeling domain is depicted in Fig. 1.

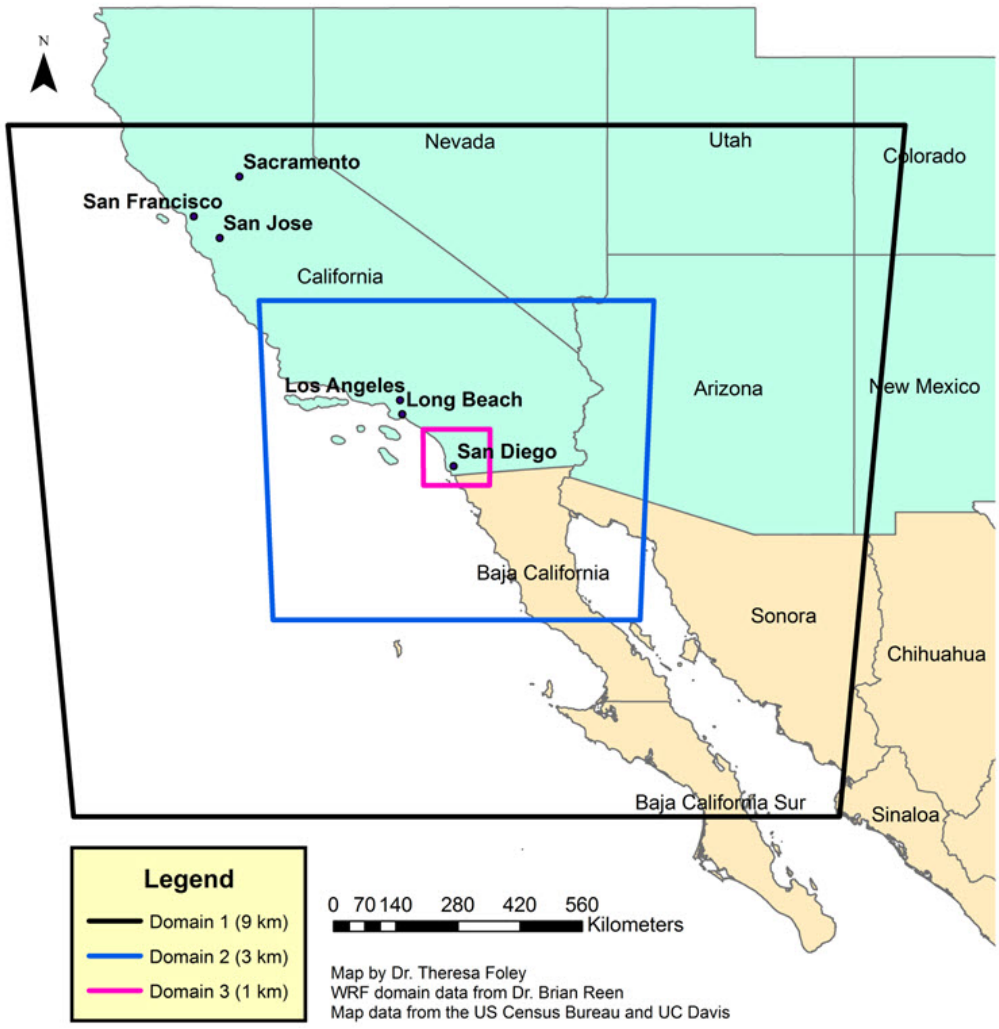

Fig. 1 Triple-nested model configuration

Note: The outer domain is dimensioned as 1,566 km by 1,566 km at a grid spacing of 9 km; a middle nest domain dimensioned as 720 km by 720 km at a grid spacing of 3 km; and an inner domain dimensioned as 126 km by 126 km at a grid spacing of 1 km. Each domain center is coincident and the nested configuration is centered near San Diego, California.

2.1 Observations

The initial conditions were constructed by starting with the GFS data as the first guess for an analysis using observations. Most observations were obtained from the Meteorological Assimilation Data Ingest System (MADIS) (National Oceanic and Atmospheric Administration [NOAA] 2014a), except for the Tropospheric Airborne Meteorological Data Reporting (TAMDAR) (Daniels et al. 2004) observations, which were obtained from AirDat (2014). The MADIS database included standard surface observations, mesonet surface observations, maritime surface observations, wind profiler measurements, rawinsonde soundings, and

Aircraft Communications Addressing and Reporting System (ACARS) data. Use and reject lists were obtained from developers of the Real-Time Mesoscale Analysis system (NOAA 2014b), and these were used to filter MADIS mesonet observations. This quality assurance evaluation is especially important given the greater tendency of mesonet observations to be poorly sited than other, more standard, surface observations.

The Obsgrid component of WRF was used for quality control of all observations. This included gross error checks, comparing observations to a background field (here GFS), and comparing observations to nearby observations. We modified Obsgrid to allow for single-level observations such as the TAMDAR and ACARS data to be more effectively compared against the GFS background field. We employed observation nudging to the observations from these same sources from 12 to 18 UTC, followed by 1-h ramping down of the nudging from 18 to 19 UTC during which time no new observations are assimilated. The forecast period thus begins at 18 UTC because no observations after this time are assimilated.

2.2 Parameterizations

For the parameterization of turbulence in WRE–N, a modified version of the Mellor–Yamada–Janjić (MYJ) Planetary Boundary Layer (Janjić 1994) scheme was used. This modification decreases the background turbulent kinetic energy and alters the diagnosis of the boundary layer depth used for model output and data assimilation (Reen et al. 2014). The WRF single-moment, 5-class microphysics parameterization is used on all domains (Hong et al. 2004), while the Kain-Fritsch (Kain 2004) cumulus parameterization is used only on the 9-km outer domain. For radiation, the Rapid Radiative Transfer Model (RRTM) parameterization (Mlawer et al. 1997) is used for long wave radiation and the Dudhia (1989) scheme for shortwave radiation. The Noah land surface model (Chen and Dudhia 2001a; 2001b) is used. Additional references and other details for these parameterization schemes are available from Skamarock et al. (2008).

2.3 Case Study Days

The case study days were selected on the basis of the prevailing synoptic weather conditions over the nested domains, and a short description of these conditions is provided in Table 1.

Table 1 Synoptic conditions for the case study days considered

Case	Dates (all 2012)	Description
1	February 07–08	Upper-level trough moved onshore, which led to widespread precipitation in the region.
2	February 09–10	Quiescent weather was in place with a 500-hPa ridge centered over central California at 12 UTC.
3	February 16–17	An upper-level low located near the California/Arizona border with Mexico at 12 UTC brought precipitation to that portion of the domain. This pattern moved south and east over the course of the day.
4	March 01–02	A weak shortwave trough resulted in precipitation in northern California at the beginning of the period that spread to Nevada, then moved southward and decreased in coverage.
5	March 05–06	Widespread high-level cloudiness due to weak upper-level low pressure but very limited precipitation.

3. Data Preparation using MET and Visual Basic

We used MET Point-Stat to compute matched pair model values for the location of each surface (2 and 10 meter) observation from the MADIS dataset in each forecast hour. To select observations closest to the top of the hour, and to eliminate multiple observations for each hour, we set the duplicate handling of Point-Stat to SINGLE. For our data, we extracted the following fields: site identification code, site latitude, site longitude, site elevation, the measured value of the meteorological variable, and the interpolated forecast value from the matched pair data for each of the 25 forecast lead times. For surface observations, MET derives the forecast value corresponding to a given observation by horizontal interpolation of the surface diagnostic variable (e.g., 2-m temperature, 2-m moisture, 10-m winds) from the enclosing grid points.

MET Point-Stat produces a single matched pair output file for each model output hour considered. Although this file is useful for many analyses, it is far easier to combine these files into a single organized structure that supports GIS analysis. We organized our data using Microsoft Access to compile the matched pair data for each forecast hour into a single database, and we used custom Visual Basic for Applications scripts to organize the data into a tabular structure that can be easily

used for GIS analysis. Generically, this table consists of 4 columns that contain site information, a column containing the meteorological variable of interest, and then a sequence of columns containing the forecast, observed, and model error (modeled–observed) for each of the 25 model output times.

During our efforts to organize our data, we discovered that several observation stations from different locations had the same site identification code. Because the differences were as much as a kilometer apart, we eliminated these station measurements from our analysis. We further restricted our analysis to the stations that reported observations in each of the 25 h of a given case day (12 UTC to 12 UTC the next day). Our intent in making this restriction was twofold: 1) we could make hour-to-hour comparisons using the same base set of data, and 2) any remaining stations within a given hour would serve as means to cross-validate our analysis. The number of stations considered for each case day is shown in Table 2, and the locations of these stations are show in Figs. 2a–e, where a–e correspond to cases 1–5.

Table 2 Count of stations that were used in each case study day, for measurements made 2 meters and 10 meters above ground level (AGL)

	Altitude (AGL)	
Case	**2 meters**	**10 meters**
1	135	72
2	96	48
3	122	58
4	133	80
5	69	47

The MADIS parameters available at 2 meters AGL include temperature, dew point, and relative humidity. The available 10-meter AGL parameters include the "u" and "v" components of the wind, as well as the magnitude and direction of the wind speed. For this study, we focused our attention on the 2-meter observations of relative humidity and temperature. We will consider the other surface parameters—such as dew point and wind—in subsequent studies.

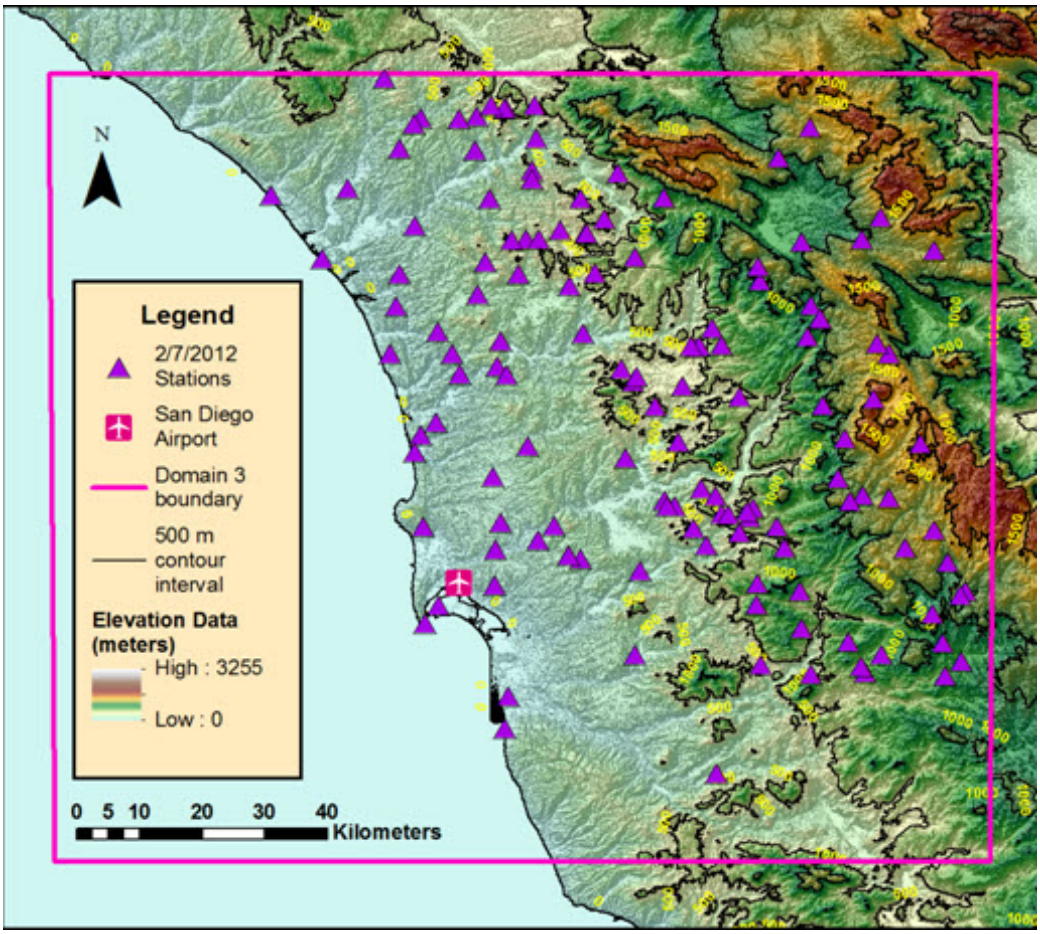

Fig. 2a The location of the weather stations that contain data for all 25 h of case day 1 overlaid on terrain data created from a US Geological Survey (USGS) digital elevation model with a resolution of 1/3 arc-second

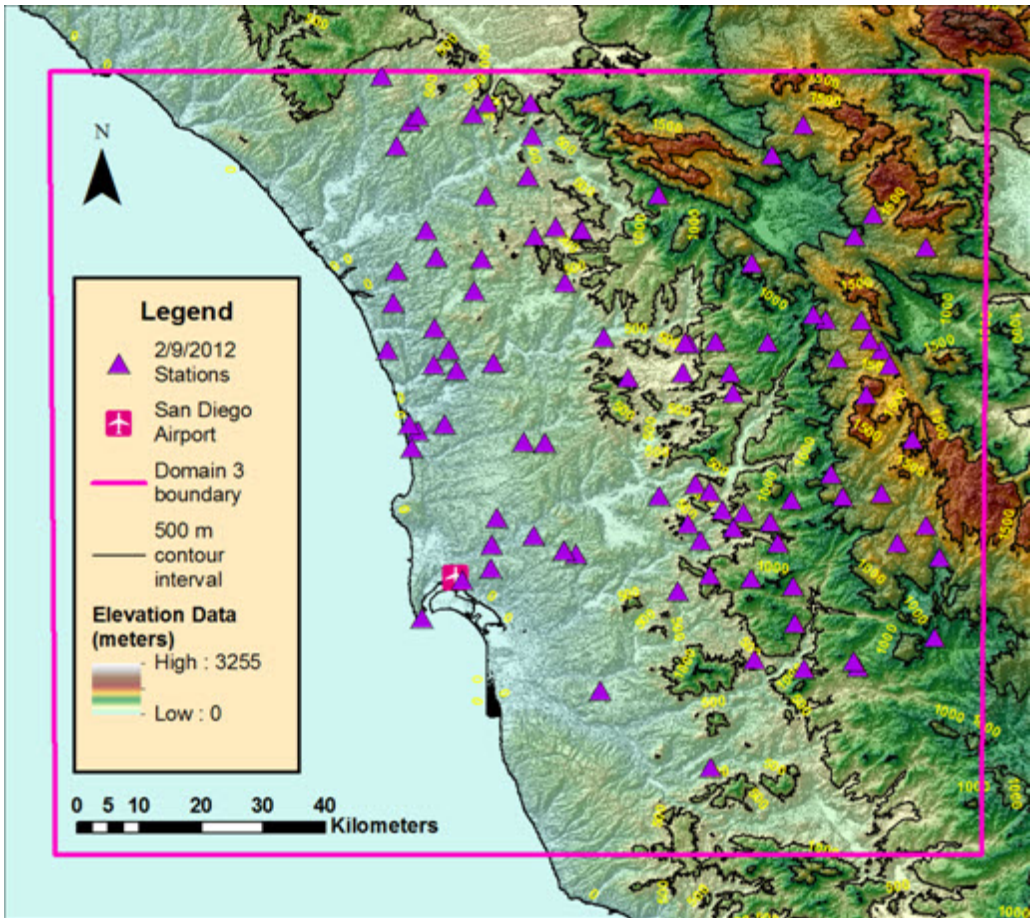

Fig. 2b The location of the weather stations that contain data for all 25 h of case day 2 overlaid on terrain data created from a USGS digital elevation model with a resolution of 1/3 arc-second

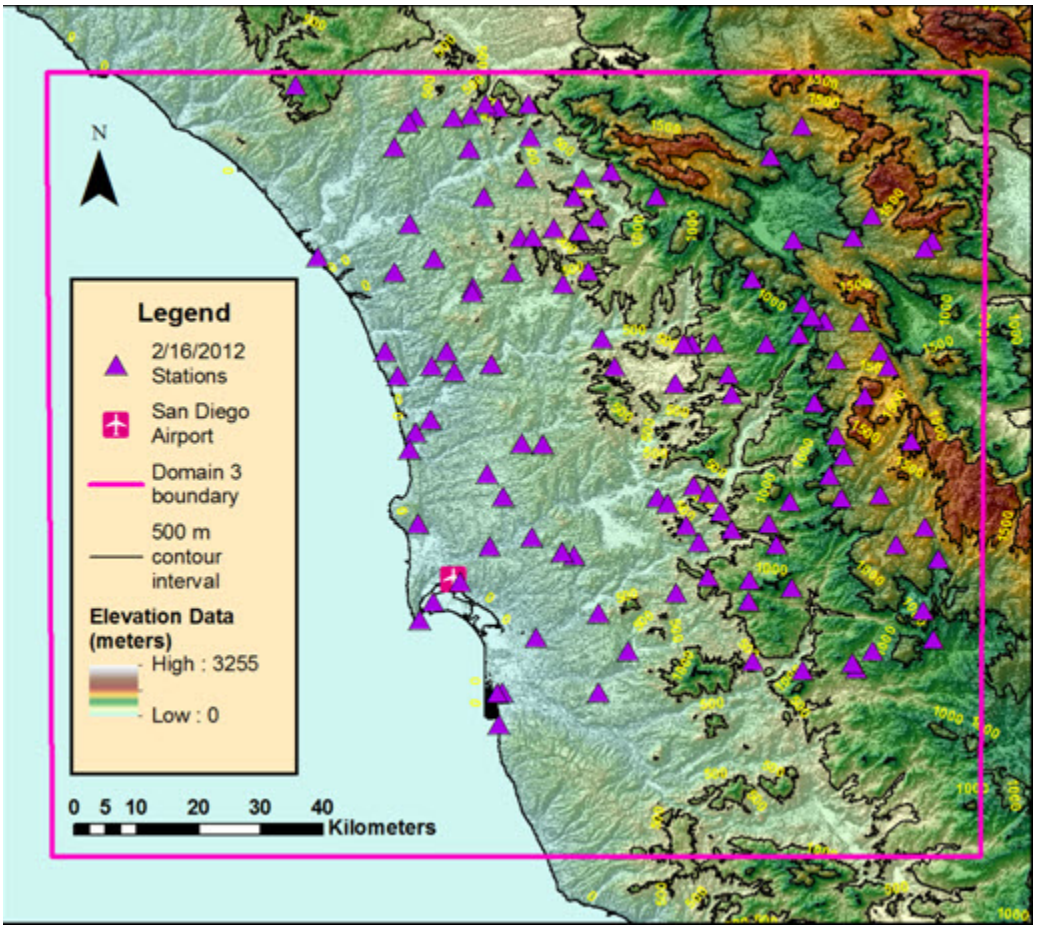

Fig. 2c The location of the weather stations that contain data for all 25 h of case day 3 overlaid on terrain data created from a USGS digital elevation model with a resolution of 1/3 arc-second

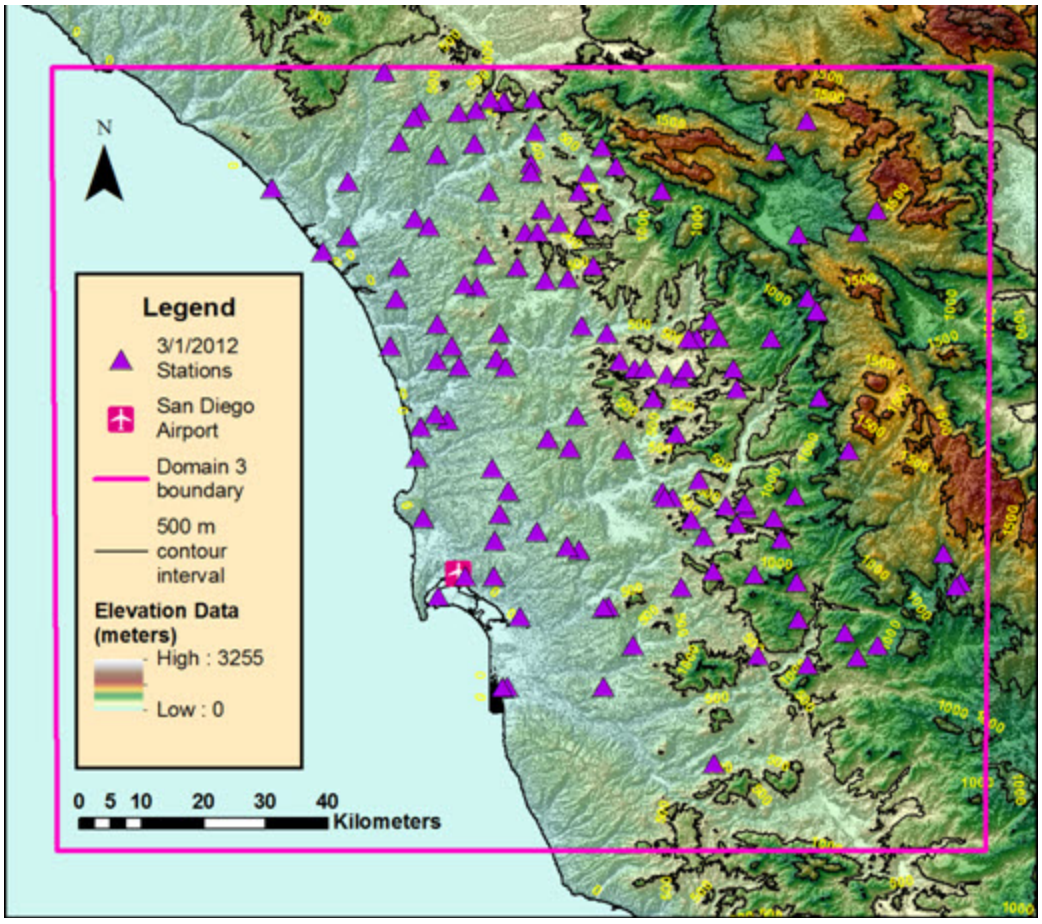

Fig. 2d The location of the weather stations that contain data for all 25 h of case day 4 overlaid on terrain data created from a USGS digital elevation model with a resolution of 1/3 arc-second

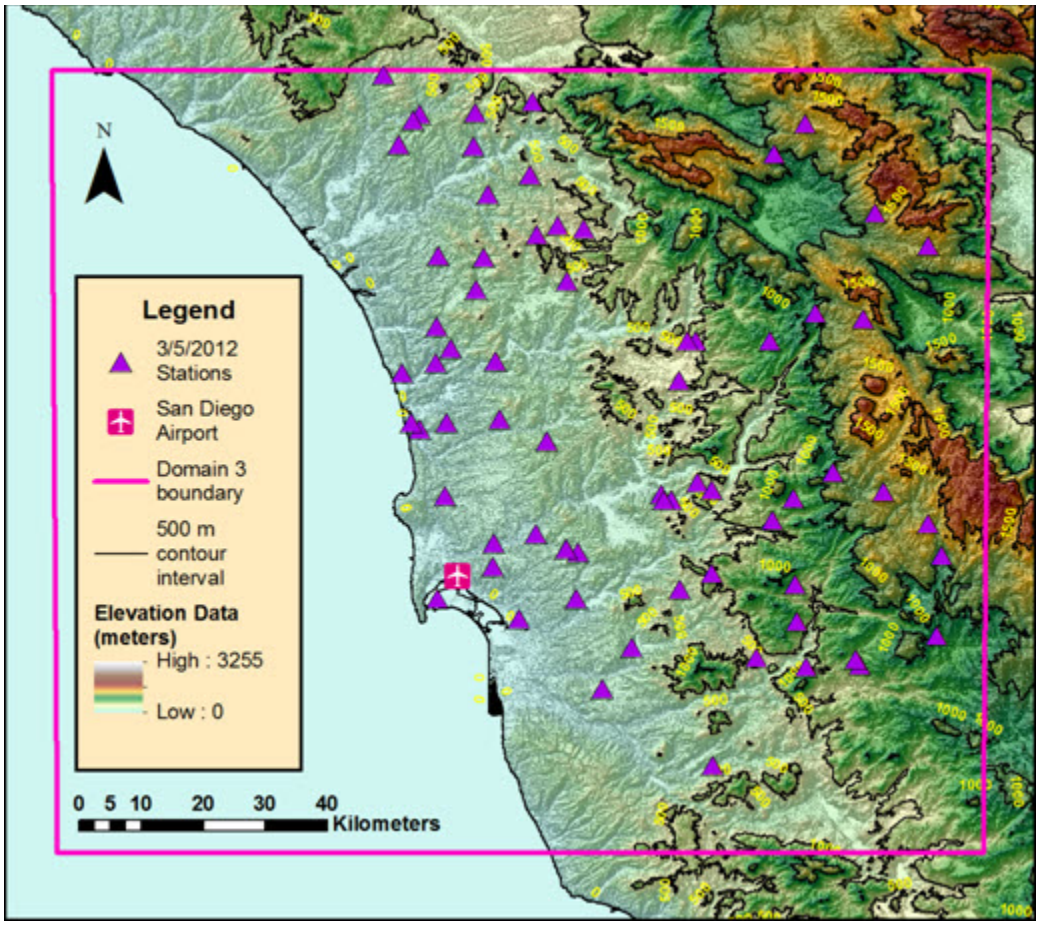

Fig. 2e The location of the weather stations that contain data for all 25 h of case day 5 overlaid on terrain data created from a USGS digital elevation model with a resolution of 1/3 arc-second

4. Data Analysis

Our first step was to examine the MET model performance statistics generated when MET created the matched-pair output data for the inner domain (domain 3). In Table 3, we summarize these statistics for 2 h in each of the 5 cases; 4:00 PM Pacific Standard Time (PST) (0000 UTC) to capture the late afternoon, unstable Planetary Boundary Layer (PBL) conditions, and 4:00 AM PST (1200 UTC) on the second day of the case to consider the early morning, stable PBL conditions. It is important to note that 4:00 AM is also the last forecast produced for that case day. Appendix A expands Table 3 to include other hours for each of the 5 case days. The average temperature bias error in Table 3 ranged from −3.55 to 2.42 K, and the average relative humidity bias error ranged from −18.66 to 17.34 K, suggesting that the model performed reasonably well over domain 3. However, because MET calculates statistics over the entire domain, large positive and

negative bias errors could be cancelling each other out and masking poor model performance.

Table 3 Mean bias (modeled–observed) errors for each of the 5 case days, and the 3 meteorological variables modeled at the Z2 surface level (2 meters AGL): 4:00 PM PST (0000 UTC) and 4:00 AM PST (1200 UTC) the next day. These statistics were calculated over the entire domain.

Case	UTC	Temperature (K)	Dew Point (K)	Relative Humidity (%)
1	0000	−0.17	1.11	5.83
	1200	0.76	1.40	2.39
2	0000	−2.89	2.95	12.27
	1200	2.42	−1.35	−13.76
3	0000	−0.65	0.09	1.94
	1200	0.81	−2.18	−11.31
4	0000	0.30	0.23	0.14
	1200	1.58	−2.18	−18.66
5	0000	−3.55	3.48	17.34
	1200	−0.13	−0.59	−1.57

Using ArcGIS version 10.2.2 (ESRI 2014), we created quantile-quantile (Q-Q) plots of the model errors (modeled value minus measured value) of each hour. A Q-Q plot provides a visual comparator of observed or experimental data to a reference probability distribution (Jolliffe and Stephenson 2011; Wilks 2011). We found that all of the model errors were approximately normally distributed; consequently, we used conventional statistical techniques that are predicated on the assumption of normality in our analysis of the model errors.

The goal of this study was to develop GIS tools for determining whether the terrain influences the model error. Previous studies have found that terrain variables such as elevation and slope have a strong influence on climatological temperature values (Agnew and Palutikof 2000; Brown and Comrie 2002; Chapman and Thornes 2003). Other candidate explanatory variables include land use (Cheng et al. 2012) and soil moisture (Massey et al. 2014). We chose elevation as the explanatory terrain variable for this study because of its known affect on temperature and conducted a correlation analysis of model error as the dependent variable against observation station elevation as the independent

variable. Future studies will include the analysis of other explanatory variables such as slope, aspect, distance from the coast, and land use.

4.1 Correlation Statistics for the Temperature and Relative Humidity Bias Errors and for Terrain Elevation

We performed a correlation analysis with model error as the dependent variable and observation elevation as the independent variable. In Figs. 3a–e, where a–e correspond to case days 1–5, we plot the Pearson correlation coefficients between the relative humidity and temperature model errors as dependent variables to elevation as the independent variable, by forecast hour for each case study day. We tested the statistical significance of each correlation coefficient using the Student t-test (Jolliffe and Stephenson 2011). Correlation coefficients of 0.2 and greater were always statistically significant with the number of measurements that were available. On case study days with a larger number of observations, some correlation coefficients below 0.2 were also statistically significant.

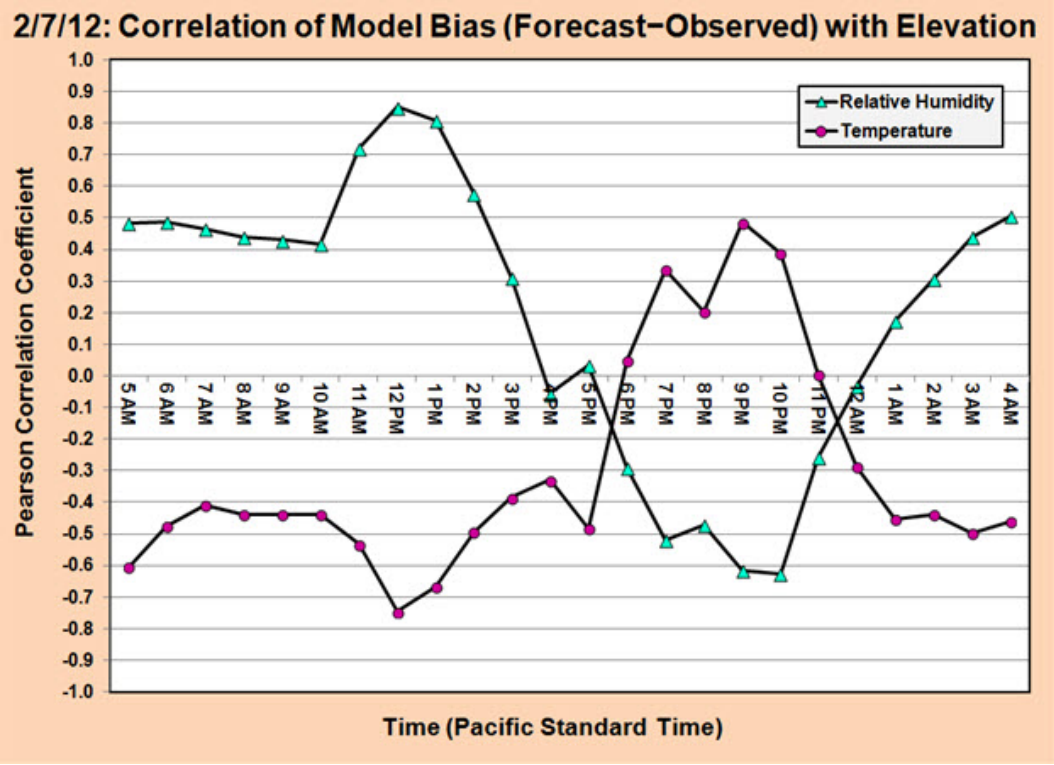

Fig. 3a The Pearson correlation coefficient computed between elevation and the model errors for relative humidity (triangles) and surface temperature (circles) for each hour of case day 1

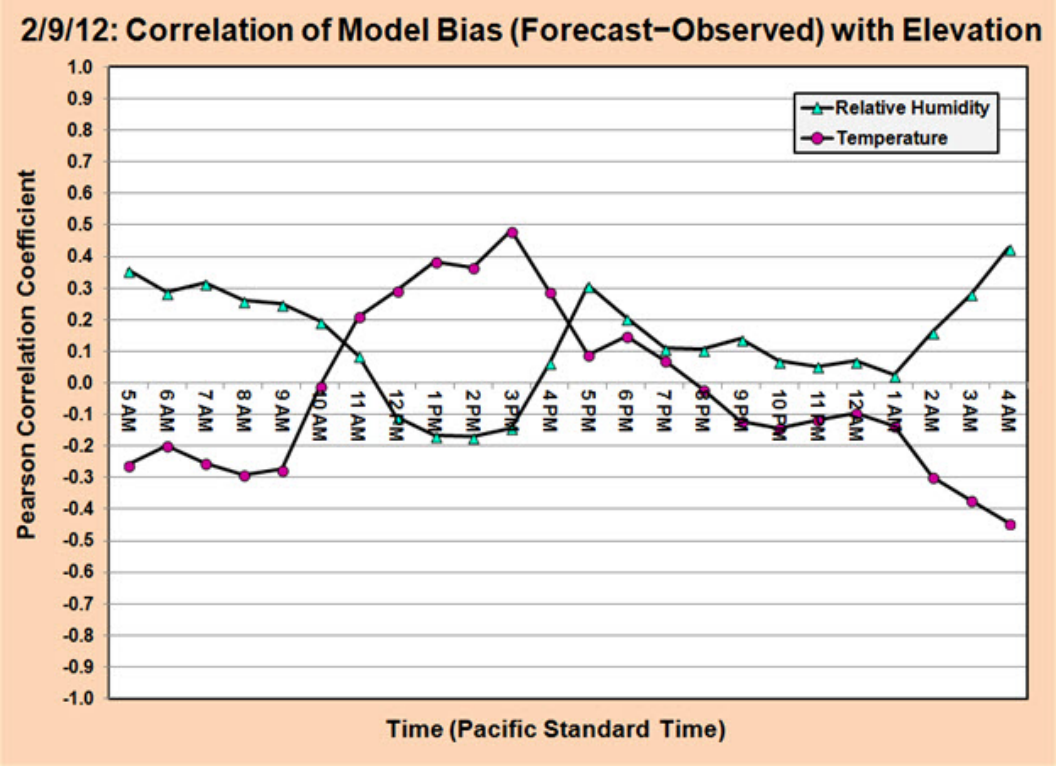

Fig. 3b The Pearson correlation coefficient computed between elevation and the model errors for relative humidity (triangles) and surface temperature (circles) for each hour of case day 2

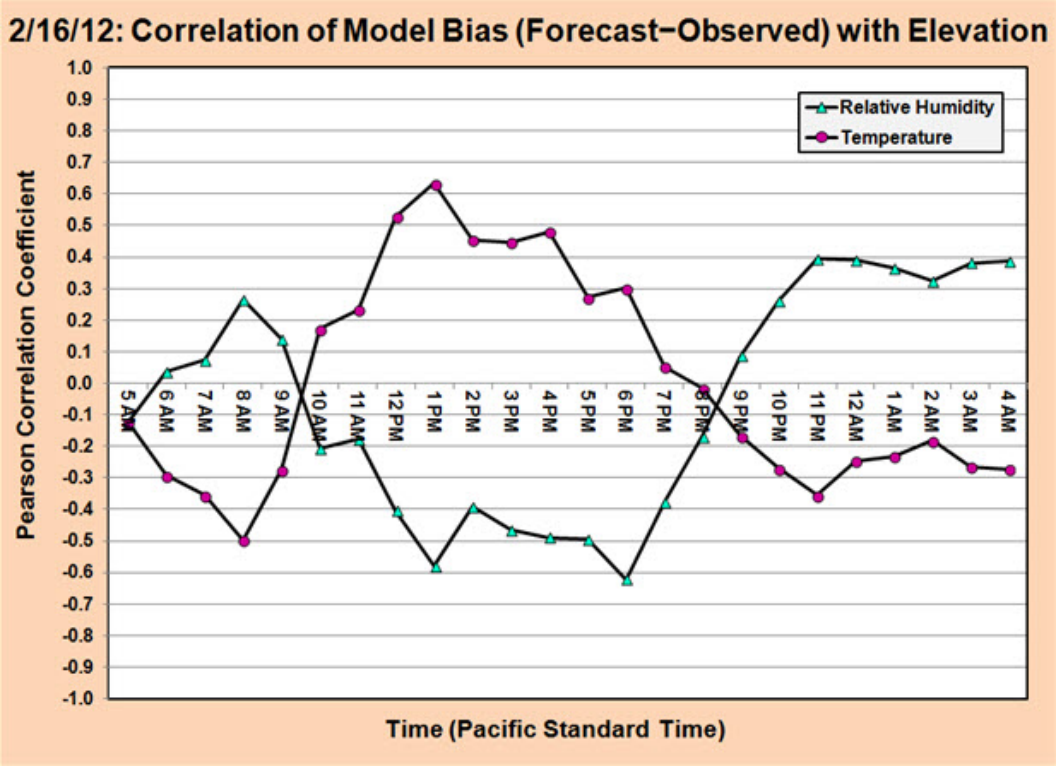

Fig. 3c The Pearson correlation coefficient computed between elevation and the model errors for relative humidity (triangles) and surface temperature (circles) for each hour of case day 3

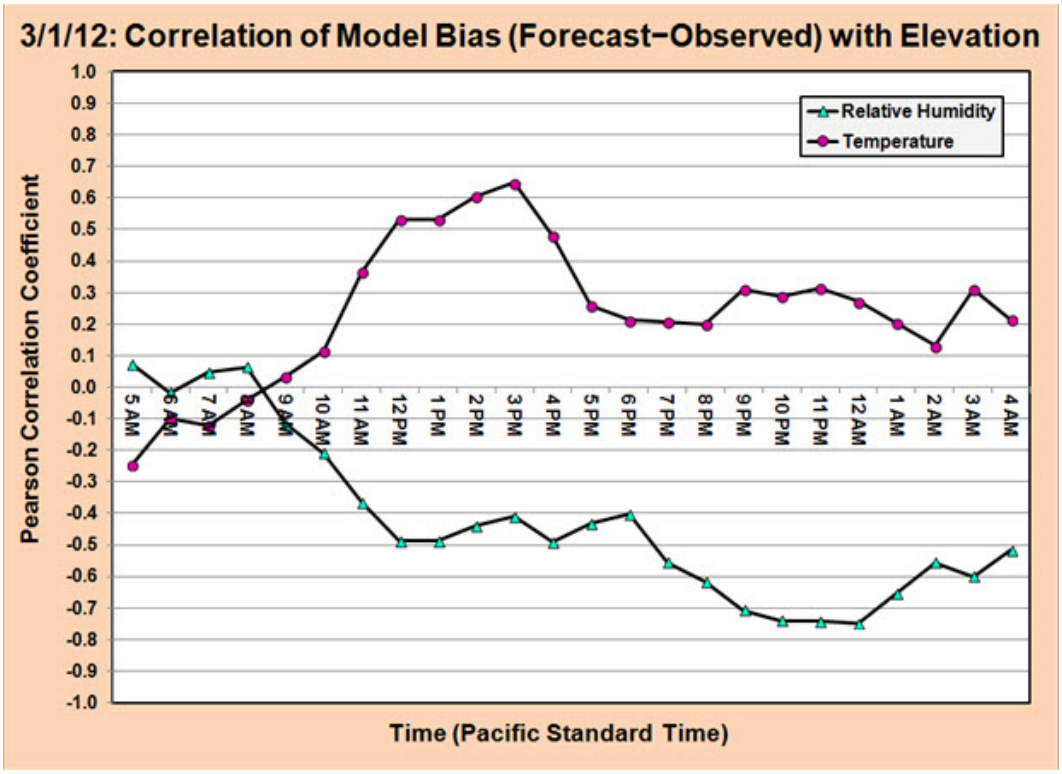

Fig. 3d The Pearson correlation coefficient computed between elevation and the model errors for relative humidity (triangles) and surface temperature (circles) for each hour of case day 4

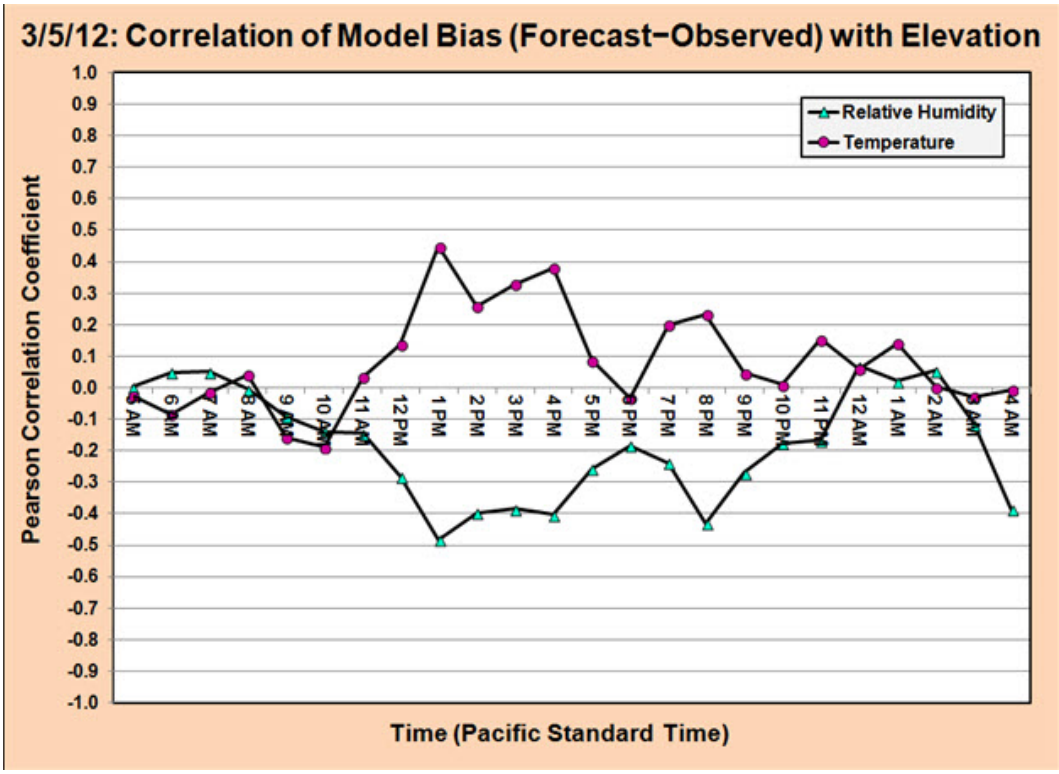

Fig. 3e The Pearson correlation coefficient computed between elevation and the model errors for relative humidity (triangles) and surface temperature (circles) for each hour of case day 5

We conducted further analysis of the data in Fig. 3 to discern diurnal trends in the correlation coefficient between temperature and relative humidity. In general, there appear to be no strong trends. However, we noted that the magnitude of the correlation of temperature with elevation to be highest in the afternoon from 12:00 PM PST to 4:00 PM PST except on February 7, the day with the most widespread precipitation. For relative humidity this was not the case. The hour of highest correlation with temperature occurred at varying times depending on the case study day.

Comparing the temporal changes in the correlation coefficients for both variables, we noticed that there appeared to be an inverse relationship between those of relative humidity and temperature. For each of the 5 case study days, we found that the correlation plots of temperature and relative humidity are nearly perfect mirror images of each other, another indication of the strong negative correlation between temperature and relative humidity forecast errors. Wherever one variable is positively correlated with elevation, the other variable is negatively correlated with elevation.

This obvious pattern raised our interest in determining why this is the case. We posed the hypothesis that the relative humidity and temperature forecast errors were correlated with each other and then calculated Pearson correlation coefficients between temperature error and relative humidity error to test this hypothesis. Figure 4 shows the plots of the coefficients by forecast hour for each case study day.

With the exception of a few hours, we found that the relative humidity and temperature error have a strong negative correlation and thus are not independent. This means that the same explanatory variables may potentially explain the variance in the forecast error of both parameters. There does not appear to be a diurnal pattern to the variation in the correlation coefficient between temperature and relative humidity, which suggests that synoptic conditions could partially account for some of the variation in Fig. 4. The most likely explanation for the strong inverse correlation between temperature and relative humidity model error lies in the underlying relationship between relative humidity and temperature. Because relative humidity is a function of the temperature and water vapor mixing ratio and an increasing temperature usually means decreasing relative humidity, a negative correlation between temperature and relative humidity error is to be expected even if the terrain is flat.

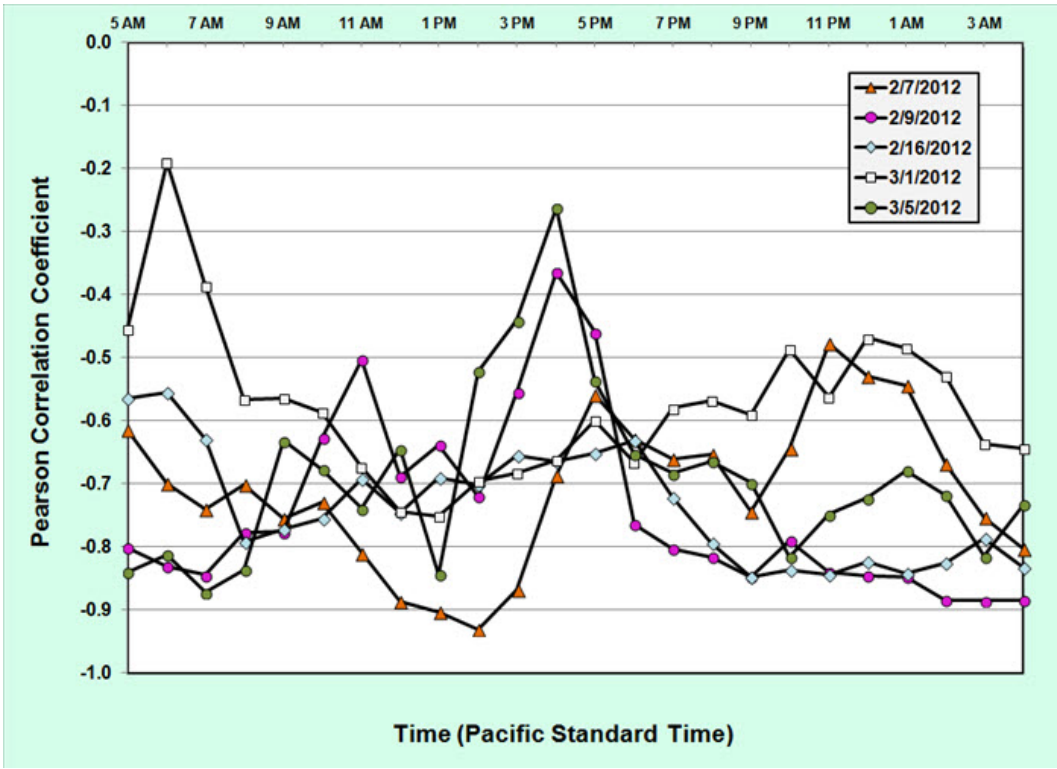

Fig. 4 The Pearson correlation coefficient between the model error of surface temperature and the model error of surface relative humidity

To quantify the contribution of terrain elevation to the overall variability of the model error, we calculated the Pearson correlation coefficients for temperature and relative humidity model error with terrain elevation. The percent of variance in the model error attributable to elevation as an explanatory variable was plotted by output hour for each case study day. Elevation accounts for a sometimes substantial portion of the variance in both the temperature and relative humidity model error for some cases and times, as shown in Figs. 5a–e, where a–e refer to case days 1–5. On average, elevation accounts for about 10% to 20% of the total variance in the model bias. It is difficult to discern a diurnal pattern, but elevation generally accounts for the greatest percentage of the model bias during the afternoon hours from 12:00 PM to 5:00 PM PST. For example, at noon on February 7, 2012 (case day 1), elevation accounts for 56% of the temperature model bias and 72% of the relative humidity model bias (elevation's strongest influence among the 5 case days investigated).

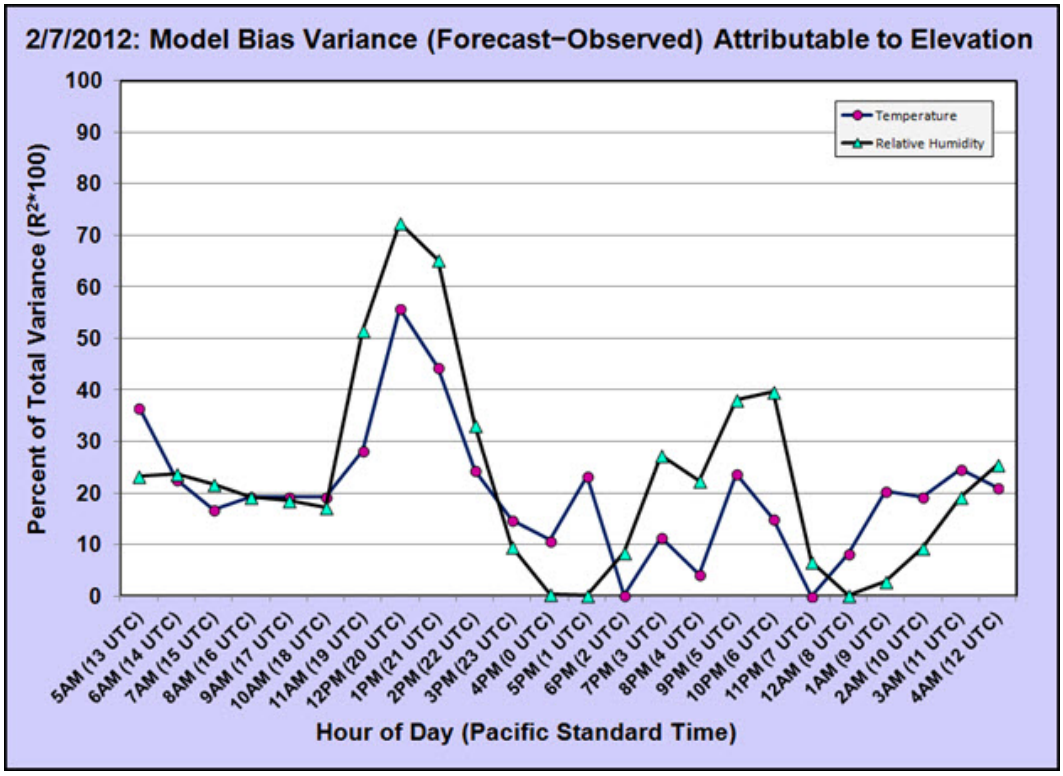

Fig. 5a The percentage of model temperature error (circles) and in the relative humidity error (triangles) variance that is attributable to elevation as an explanatory variable for case day 1

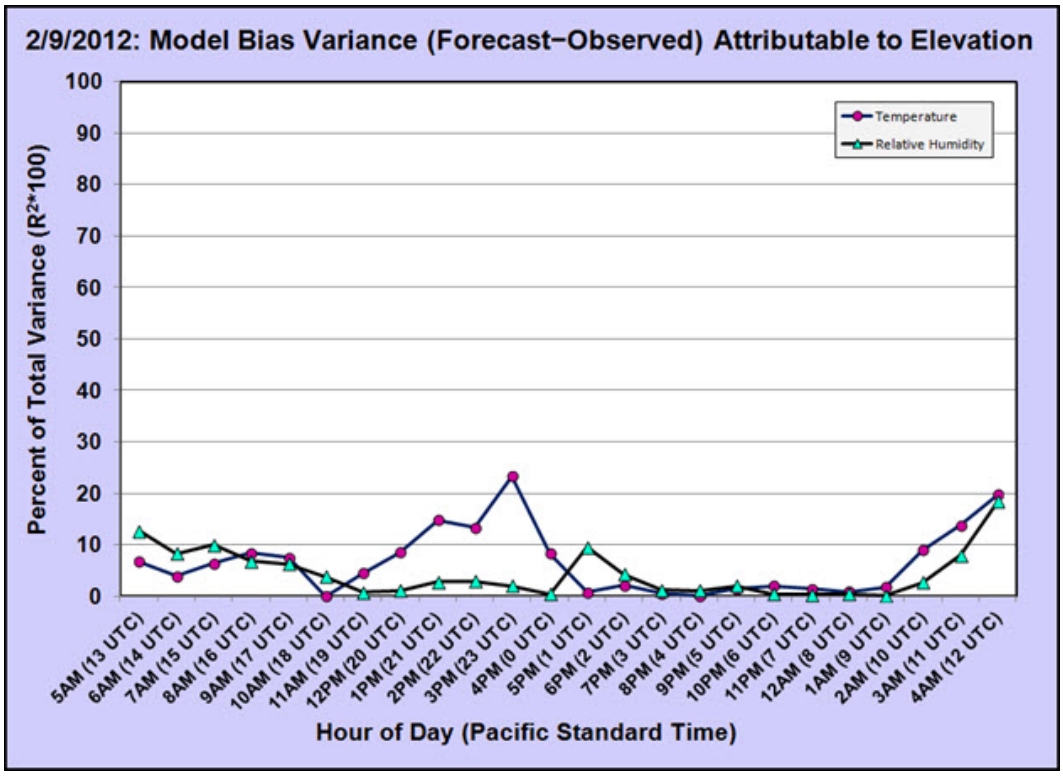

Fig. 5b The percentage of model temperature error (circles) and in the relative humidity error (triangles) variance that is attributable to elevation as an explanatory variable for case day 2

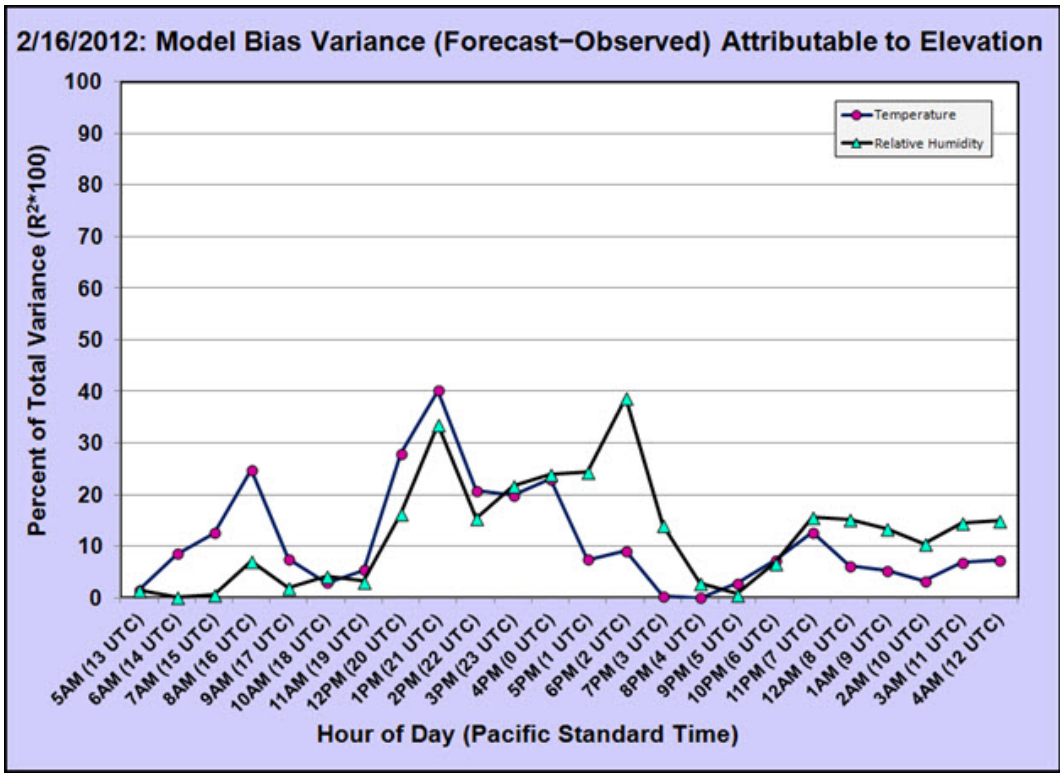

Fig. 5c The percentage of model temperature error (circles) and in the relative humidity error (triangles) variance that is attributable to elevation as an explanatory variable for case day 3

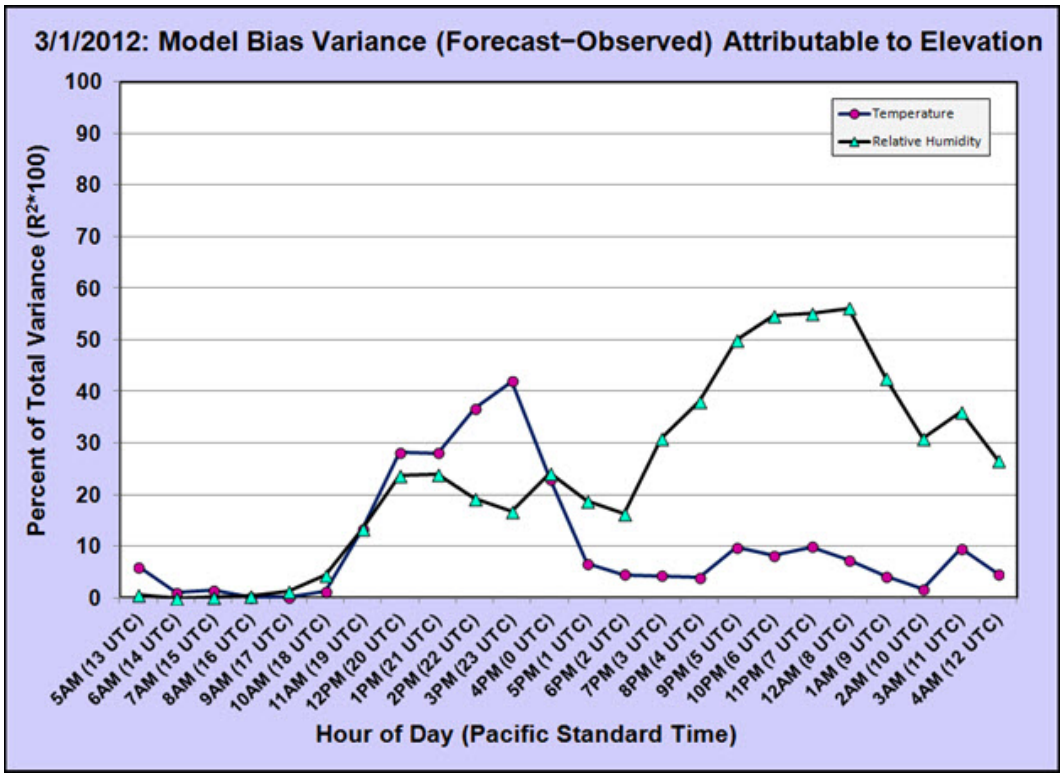

Fig. 5d The percentage of model temperature error (circles) and in the relative humidity error (triangles) variance that is attributable to elevation as an explanatory variable for case day 4

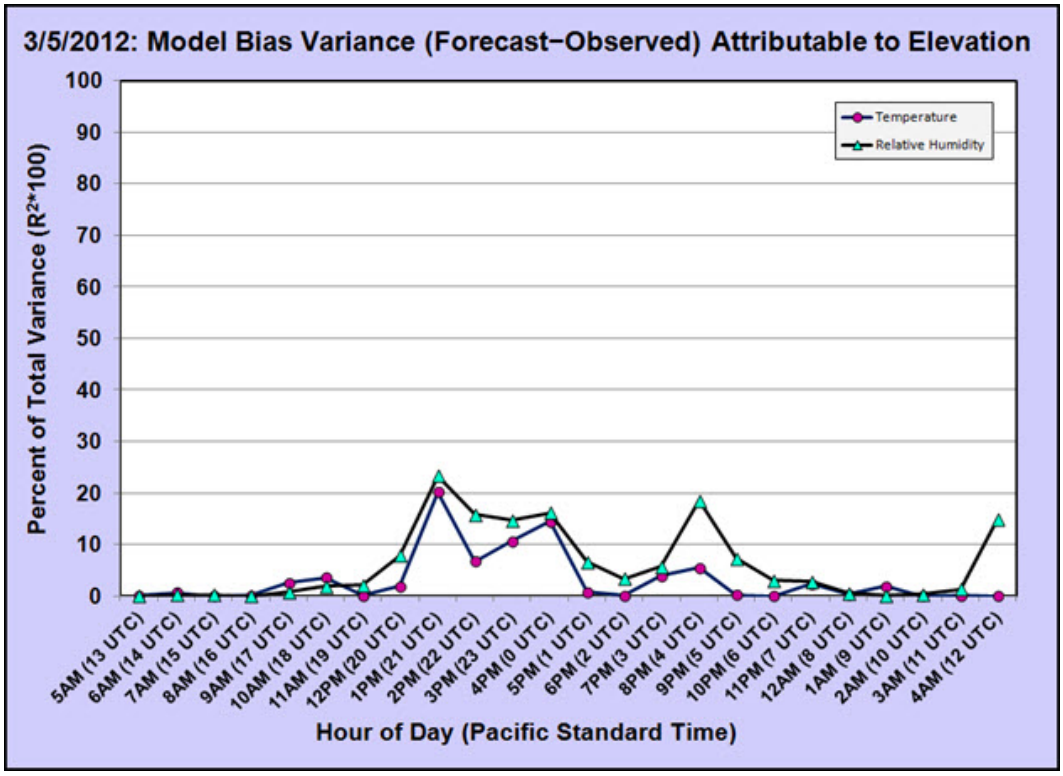

Fig. 5e The percentage of model temperature error (circles) and in the relative humidity error (triangles) variance that is attributable to elevation as an explanatory variable for case day 5

For each case day, we determined the number of model output hours out of the 25 available (6 analysis and 19 forecast hours), which had elevation as a significant explanatory variable. Because the number of observation stations considered for a given case day was fixed for that case (see Table 2), we could determine a threshold (at an α of 5%) for the correlation coefficient above which the correlation would be considered significant. Finally, we judged that if a large number (more than half) of the model output hours had significant correlations, then elevation was a significant explanatory variable for that variable and case. This data is presented in Table 4 for both Temperature and Relative Humidity.

Table 4 Statistics related to the percentage of variance of the temperature and relative humidity forecast errors explained by elevation

Note: For both threshold and maximum R^2, the values are the Pearson R^2 coefficient multiplied by 100 to show the values as percentiles.

Temperature				
Case Day	Threshold R^2 (%)	No. Significant Output Hours	Maximum R^2 (%)	Elevation Significant?
1	4.19	22	55.71	Yes
2	3.99	14	23.29	Yes
3	2.79	21	40.18	Yes
4	4.01	18	41.98	Yes
5	5.43	5	20.18	No

Relative Humidity				
Case Day	Threshold R^2 (%)	No. Significant Output Hours	Maximum R^2 (%)	Elevation Significant?
1	2.95	21	72.30	Yes
2	4.26	9	18.32	No
3	2.84	19	38.68	Yes
4	4.37	19	56.14	Yes
5	5.71	10	23.29	No

4.2 Estimating the Forecast Error over the Entire Domain using GIS

Using a GIS, we created a raster mosaic of terrain data from USGS digital elevation model (DEM) data (USGS 2013). The USGS DEM has a resolution of 1/3 arc-second and a datum of NAD1983. A datum is a mathematical algorithm to account for the fact that the earth is not perfectly spherical. A projection corrects for the distortion that occurs when projecting spherical latitude and longitude coordinates onto the flat surface of paper or a computer screen. WRF uses no datum (assumes that the earth is a perfect sphere), and we used the Lambert Conic Conformal projection in our WRF modeling.

Matching the Point-Stat matched pair data with the DEM proved to be a big challenge, because the projection and datum of the MADIS data were not specified. Upon the advice of Olga Wilhelmi and Jennifer Boehnert from NCAR (2014), we used the NAD1983 datum without a projection for the Point-Stat data. With these GIS settings, the Point-Stat data matched the USGS DEM very well.

For example, weather stations located on the tiny Channel Islands off the California coast were correctly geo-located on land rather than the Pacific Ocean.

To estimate the model errors over the entire domain, we treated each weather station measurement as a sample point from a random field and employed the Empirical Bayesian Kriging (EBK) tool (Krivoruchko 2011; 2012) of ArcGIS to estimate a continuous error surface. These surfaces do not extend to the western or southern extents of the domain because there is insufficient data in these areas to produce a statistically significant surface estimate. Contours of the bias error, corresponding to the quintiles of the error distribution, were created to ensure that the underlying terrain was visible.

In our analysis of the 5 case study days, we studied the 5 specific hours given in Table 5. We chose these specific times to examine the impact of the data assimilation and the stability of the boundary layer. Under clear skies and weak synoptic forcing, at 5:00 AM PST (the second assimilation hour) we would expect the boundary layer to be stable because the sun has not yet warmed the surface of the earth. Similarly the boundary layer is often unstable at about 0000 UTC due to a continual upward heat flux following the earlier maximum solar insolation period. As the ground cools after sunset, the boundary layer becomes more stable. Clear sky conditions were less common during the 5 case study days of this study. Precipitation occurred over the study domain on case days 1, 3, and 4; and much of the domain was cloudy on case day 5. Under cloudy conditions with precipitation, we expect that the actual tendency would have been for more neutral/less stable conditions during nighttime and less unstable during the day.

Table 5 Hours chosen for further GIS study and the rationale for these choices

	Time (h)			
Period	**UTC**	**Forecast Lead**	**Local (PST)**	**Reason**
Analysis	13	–5	5:00 AM	First hour of data assimilation completed
	18	0	10:00 AM	Beginning of forecast; near neutral boundary layer
Forecast	0	6	4:00 PM	Late afternoon, unstable boundary layer
	6	12	10:00 PM	Early evening, slightly stable boundary layer
	12	18	4:00 AM	Early morning, stable boundary layer

The effect of the terrain on the model error can be clearly seen in the March 1, 2012 (case day 4) error contour maps of Fig. 6. For most of the day, the temperature model bias is positively correlated with elevation, while relative humidity model bias is negatively correlated with elevation. However, an

exception is 5:00 AM PST (13 UTC; Figure 6a, top row), when these correlations are reversed; however, the correlation coefficients for this hour are not statistically significant. The strong inverse relationship between temperature and relative humidity is clearly evident in the afternoon hours of March 1 (Fig. 6b, top row). The terrain effects can also be seen in the forecast error contour maps for February 9, February 26, and March 5, 2012, which can be found in Appendix B.

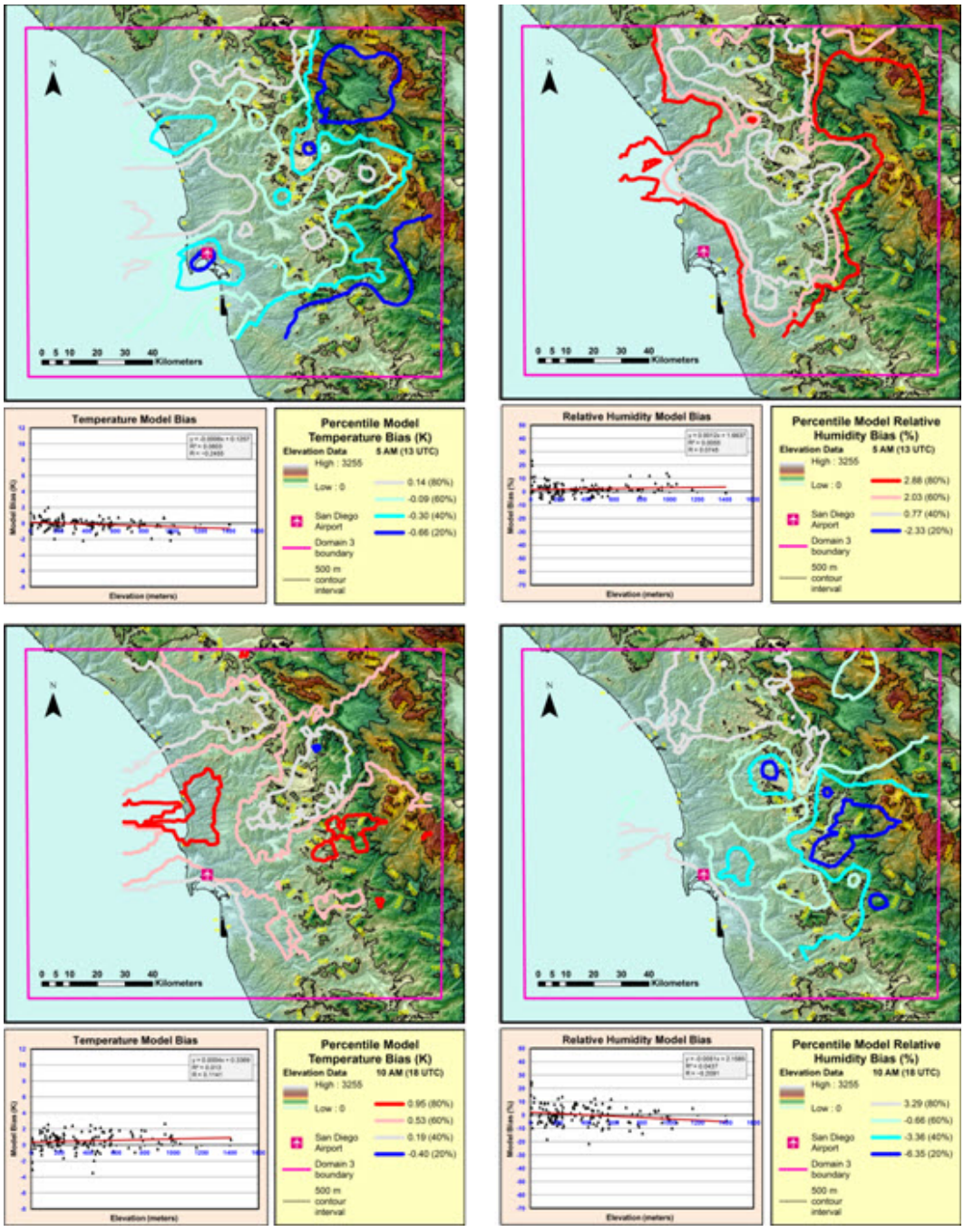

Fig. 6a EBK-created isolines of constant model error (forecast minus observed) for March 1, 2012 (case day 4), for 5:00 AM PST (13 UTC) and 10:00 AM PST (18 UTC)

Note: The temperature model error is shown in the left column and the relative humidity model error in the right column. Each map shows USGS terrain height (shaded) overlaid with colored isolines showing the quintiles of the error distribution. The scatterplot with each map shows the relationship between terrain height and model error along with a best fit line. The top row is 5:00 AM PST (13 UTC, 1 h into preforecast) and the bottom row is 10:00 AM PST (18 UTC, 0-h forecast).

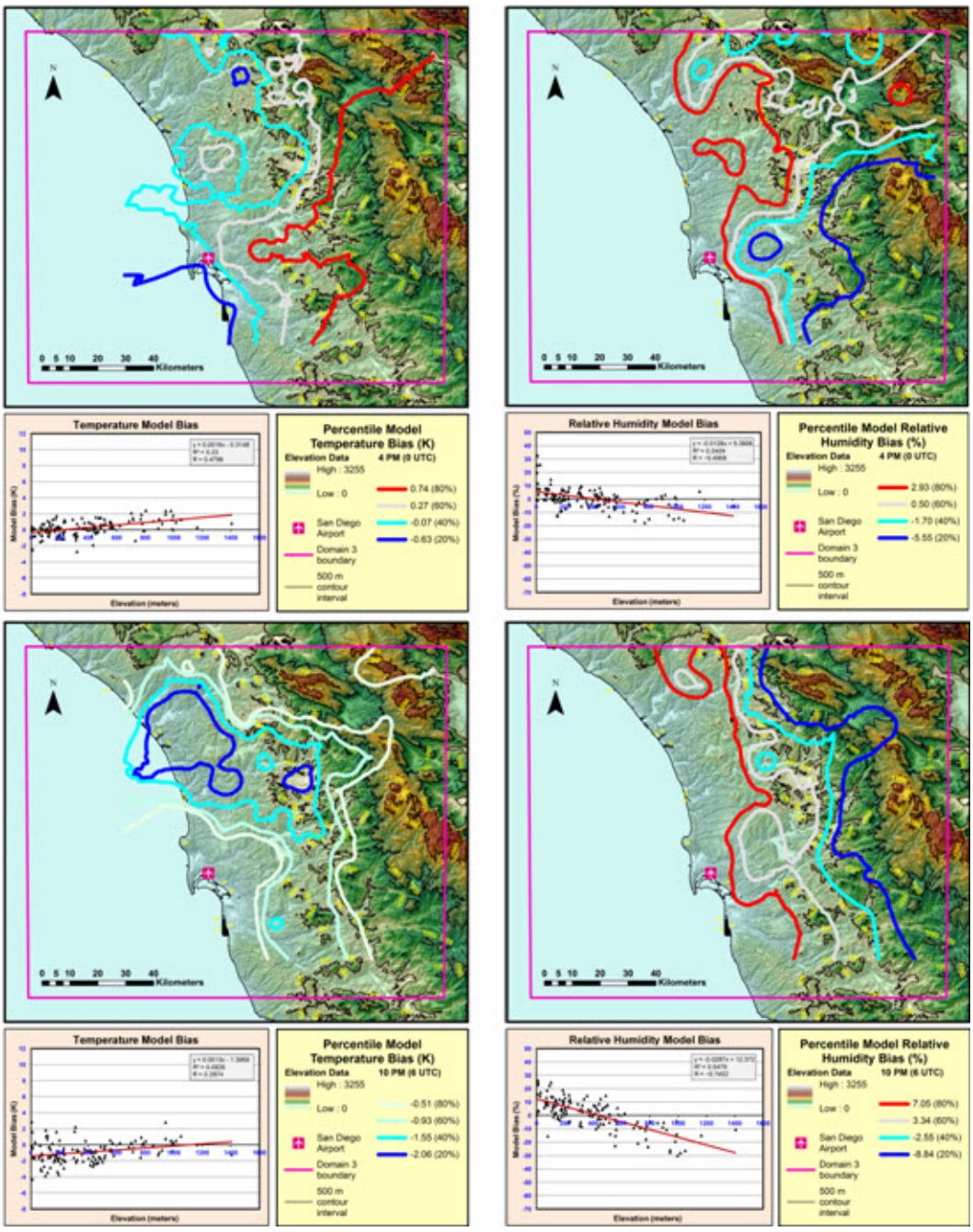

Fig. 6b EBK-created isolines of constant model error (forecast minus observed) for March 1, 2012 (case day 4), for 4:00 PM PST (0 UTC) and 10:00 PM PST (6 UTC)

Note: The temperature model error is shown in the left column and the relative humidity model error in the right column. Each map shows USGS terrain height (shaded) overlaid with colored isolines showing the quintiles of the error distribution. The scatterplot with each map shows the relationship between terrain height and model error along with a best fit line. The top row is 4:00 PM PST (0 UTC, 6-h forecast) and the bottom row is 10:00 PM PST (6 UTC, 12-h forecast).

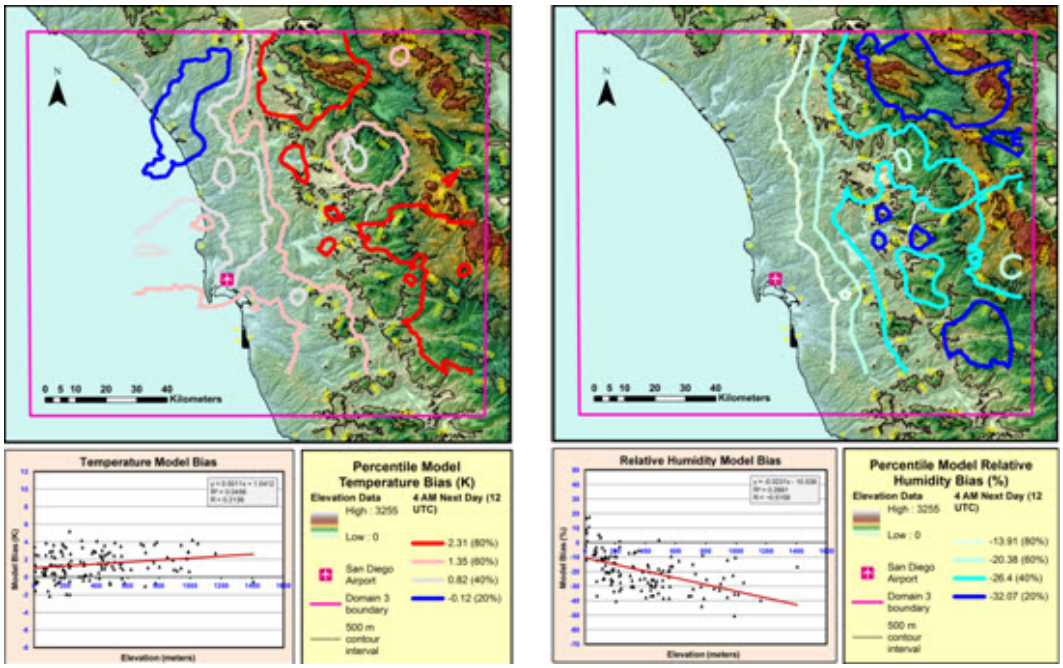

Fig. 6c EBK-created isolines of constant model error (forecast minus observed) for March 1, 2012 (case day 4), for 4:00 AM PST (12 UTC) on March 2, 2012

Note: The temperature model error is shown in the left column and the relative humidity model error in the right column. Each map shows USGS terrain height (shaded) overlaid with colored isolines showing the quintiles of the error distribution. The scatterplot with each map shows the relationship between terrain height and model error along with a best fit line. This row is for 4:00 AM PST the next day (12 UTC, 18-h forecast).

We also created error contours for each hour of February 7, 2012. Selected hours are shown in Figs. 7a–g, and error contours for every hour can be found in Appendix C. From 10:00 AM PST to 1:00 PM PST (Figs. 7a and 7b, respectively) temperature model error is negatively correlated with elevation, while the relative humidity model error is positively correlated. These correlations are evident in both the error contours and the scatter plots of forecast error versus elevation. Between 5:00 PM and 6:00 PM PST (Fig. 7c), these correlations are reversed so that temperature forecast error is positively correlated with elevation, while the relative humidity forecast error is negatively correlated. This trend continues from 7:00 PM PST to 10:00 PM PST (Figs. 7d and 7e, respectively). Between 11:00 PM and 12:00 AM PST (Fig. 7f), the correlations are again reversed.

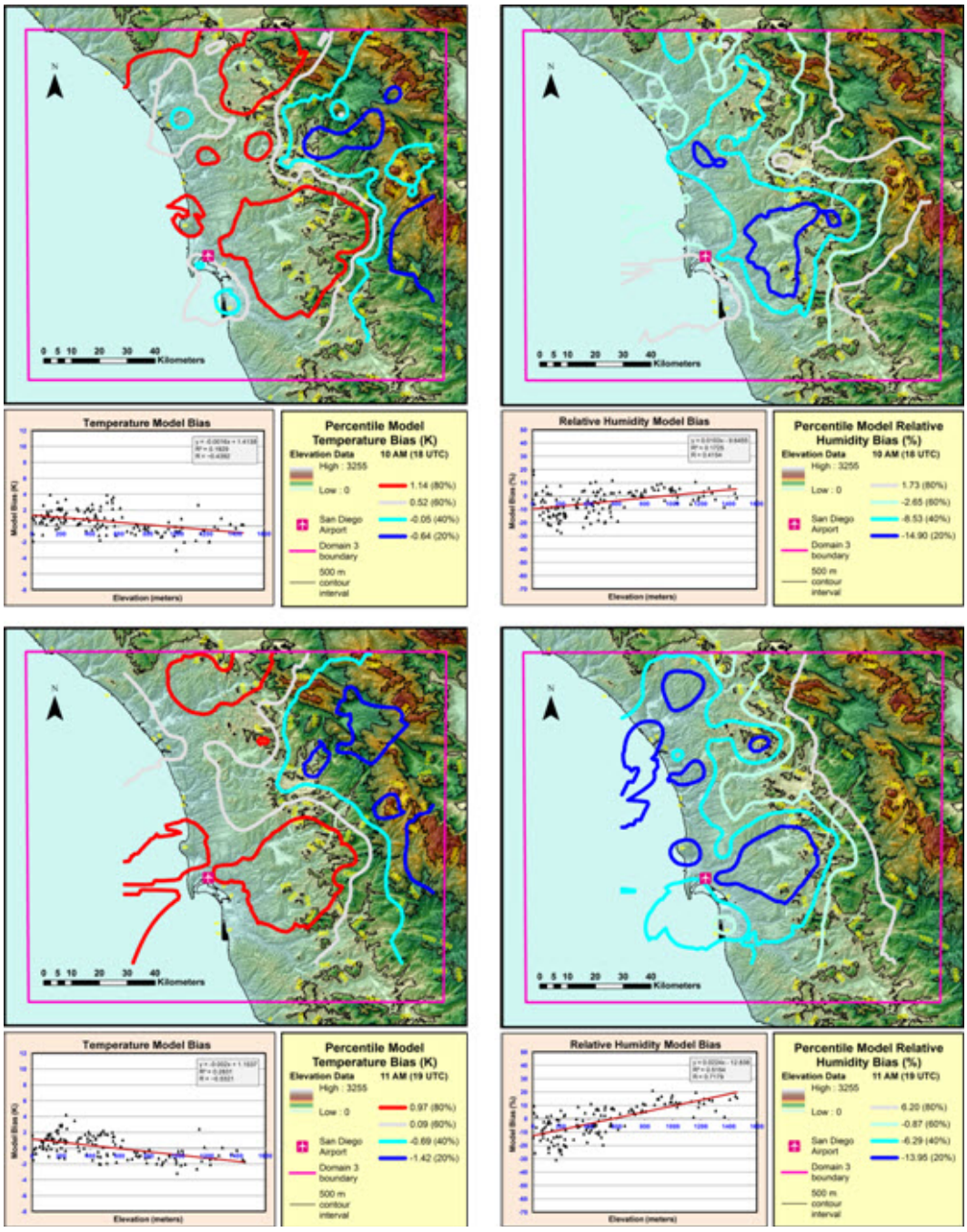

Fig. 7a EBK-created isolines of constant model error (forecast minus observed) for February 7, 2012 (case day 1) for 10:00 AM PST (18 UTC) and 11:00 AM PST (19 UTC)

Note: The temperature model error is shown in the left column and the relative humidity model error in the right column. Each map shows USGS terrain height (shaded) overlaid with colored isolines showing the quintiles of the error distribution. The scatterplot with each map shows the relationship between terrain height and model error along with a best fit line. The top row is 10:00 AM PST (18 UTC, 0-h forecast) and the bottom row is 11:00 AM PST (19 UTC, 1-h forecast).

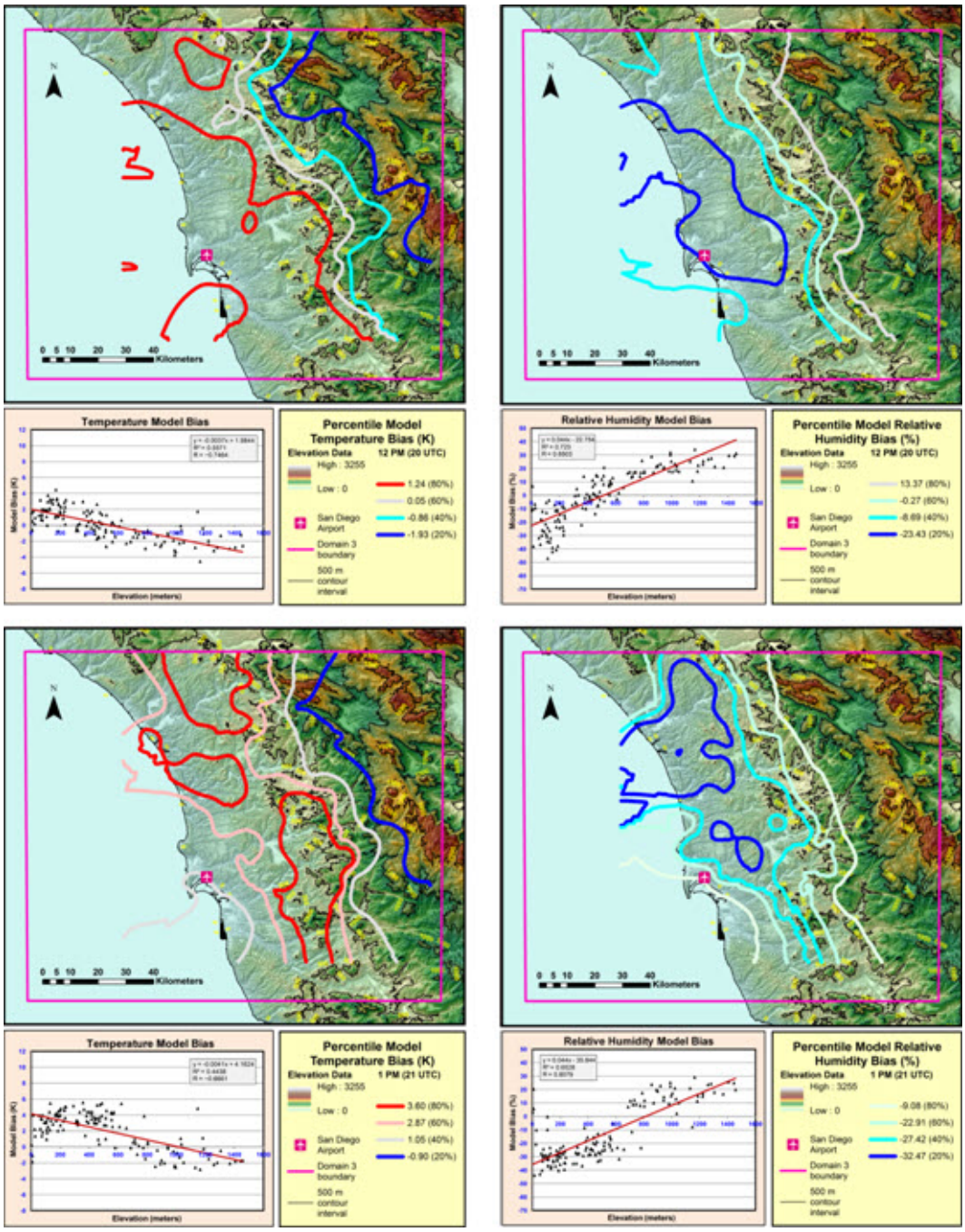

Fig. 7b EBK-created isolines of constant model error (forecast minus observed) for February 7, 2012 (case day 1) for 12:00 AM PST (20 UTC) and 1:00 AM PST (21 UTC)

Note: The temperature model error is shown in the left column and the relative humidity model error in the right column. Each map shows USGS terrain height (shaded) overlaid with colored isolines showing the quintiles of the error distribution. The scatterplot with each map shows the relationship between terrain height and model error along with a best fit line. The top row is 12:00 PM PST (20 UTC, 2-h forecast) and the bottom row is 1:00 PM PST (21 UTC, 3-h forecast).

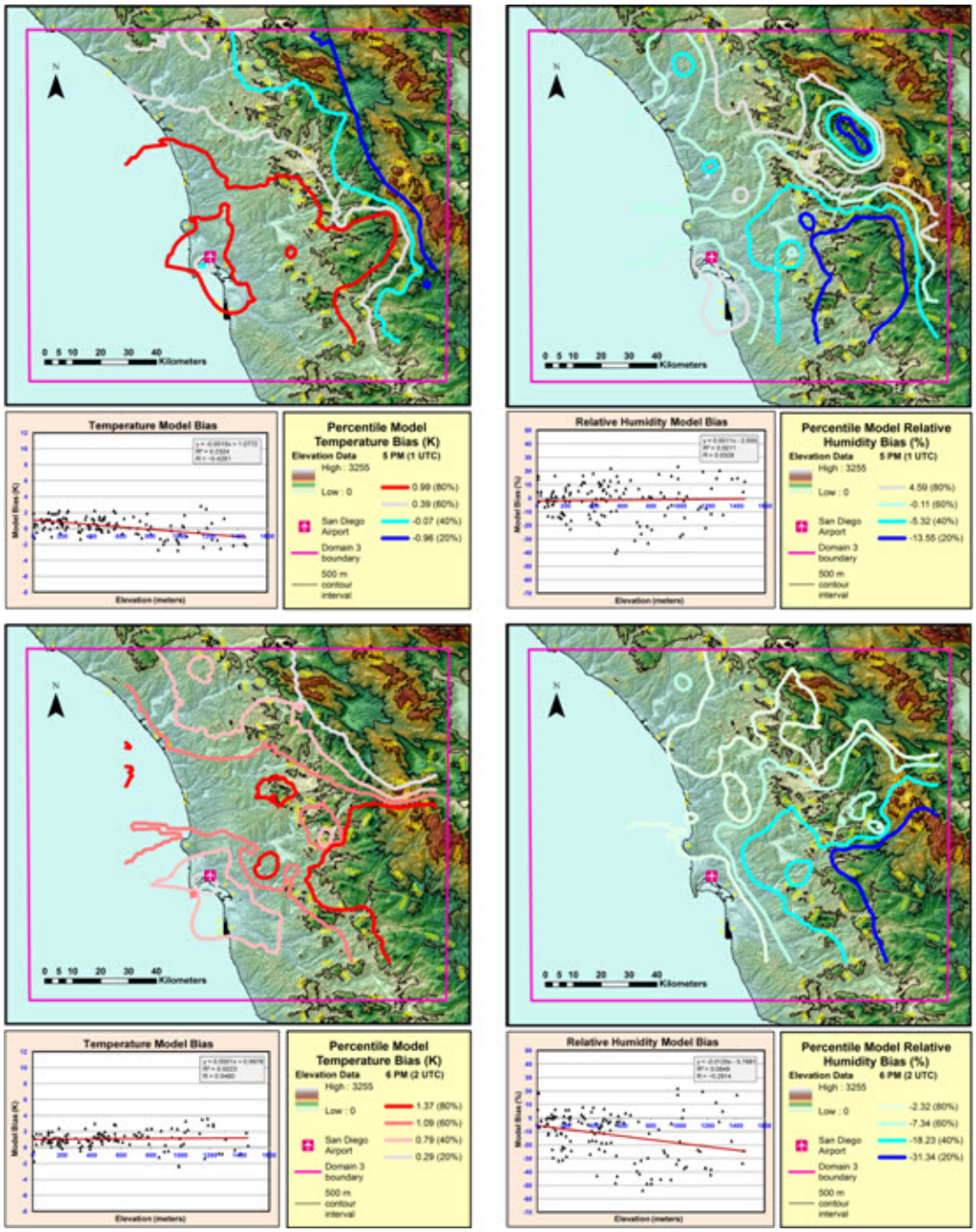

Fig. 7c EBK-created isolines of constant model error (forecast minus observed) for February 7, 2012 (case day 1) for 5:00 PM PST (1 UTC) and 6:00 PM PST (2 UTC)

Note: The temperature model error is shown in the left column and the relative humidity model error in the right column. Each map shows USGS terrain height (shaded) overlaid with colored isolines showing the quintiles of the error distribution. The scatterplot with each map shows the relationship between terrain height and model error along with a best fit line. The top row is 5:00 PM PST (1 UTC, 7-h forecast) and the bottom row is 6:00 PM PST (2 UTC, 8-h forecast).

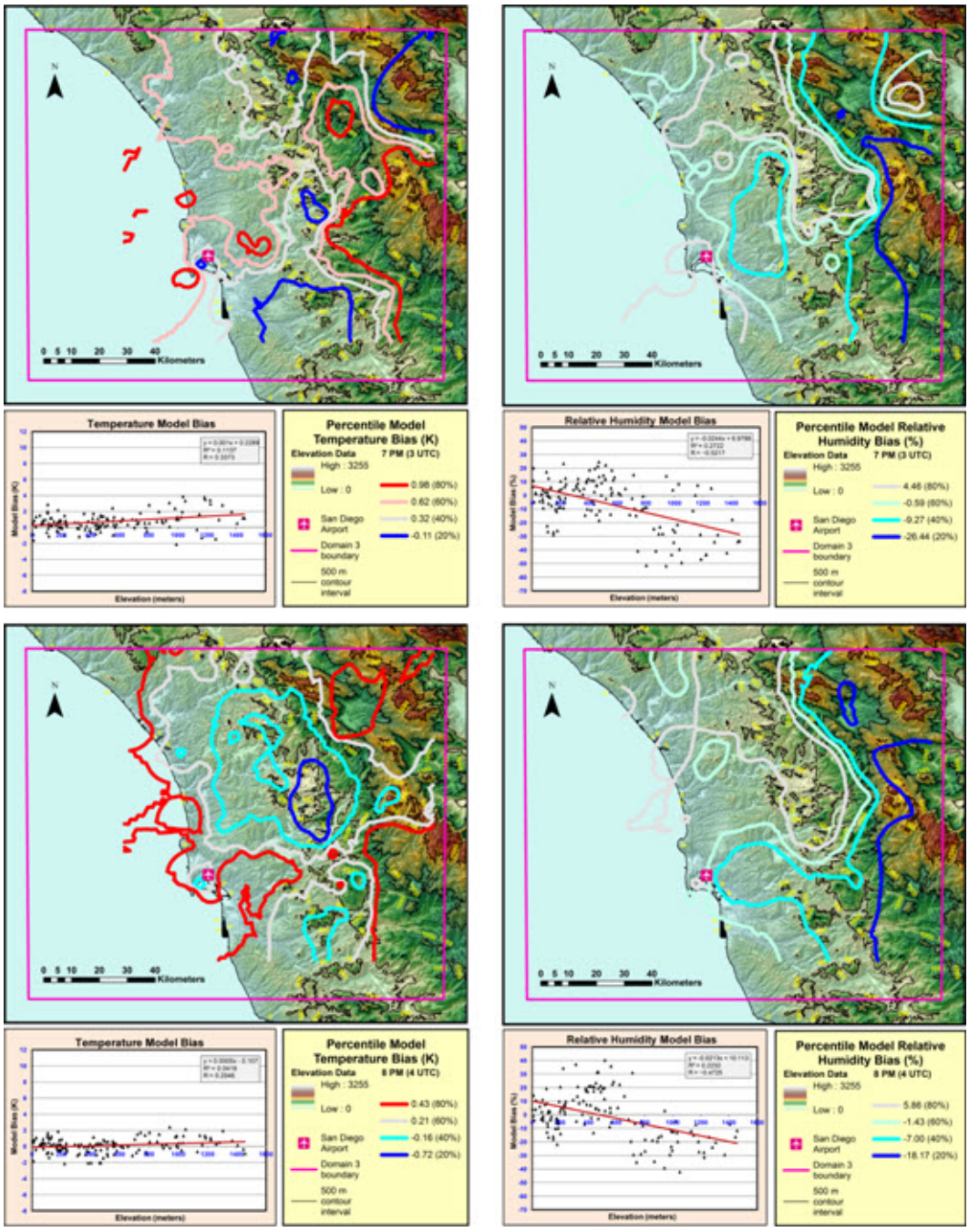

Fig. 7d EBK-created isolines of constant model error (forecast minus observed) for February 7, 2012 (case day 1) for 7:00 PM PST (3 UTC) and 8:00 PM PST (4 UTC)

Note: The temperature model error is shown in the left column and the relative humidity model error in the right column. Each map shows USGS terrain height (shaded) overlaid with colored isolines showing the quintiles of the error distribution. The scatterplot with each map shows the relationship between terrain height and model error along with a best fit line. The top row is 7:00 PM PST (3 UTC, 9-h forecast) and the bottom row is 8:00 PM PST (4 UTC, 10-h forecast).

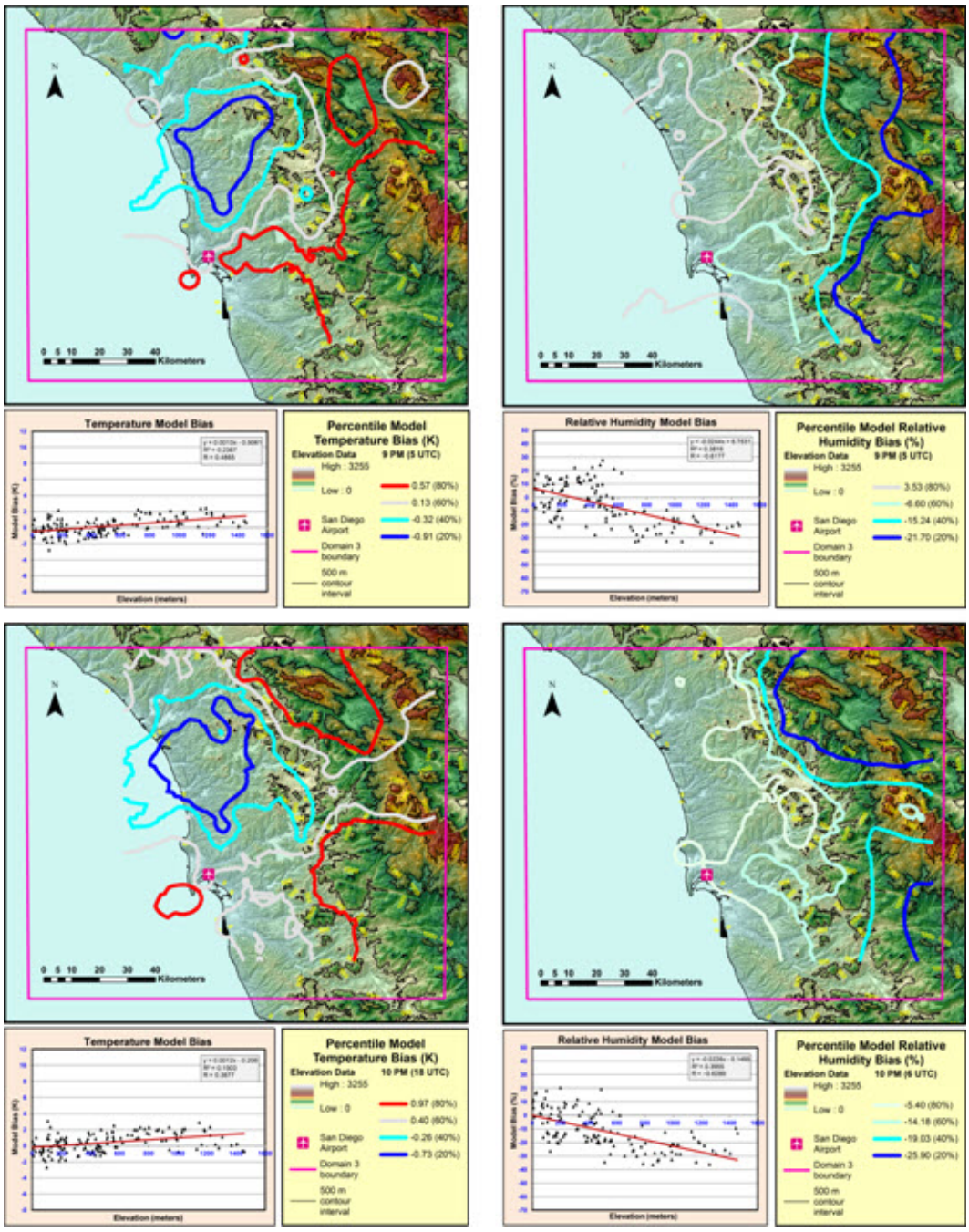

Fig. 7e EBK-created isolines of constant model error (forecast minus observed) for February 7, 2012 (case day 1) for 9:00 PM PST (5 UTC) and 10:00 PM PST (6 UTC)

Note: The temperature model error is shown in the left column and the relative humidity model error in the right column. Each map shows USGS terrain height (shaded) overlaid with colored isolines showing the quintiles of the error distribution. The scatterplot with each map shows the relationship between terrain height and model error along with a best fit line. The top row is 9:00 PM PST (5 UTC, 11-h forecast) and the bottom row is 10:00 PM PST (6 UTC, 12-h forecast).

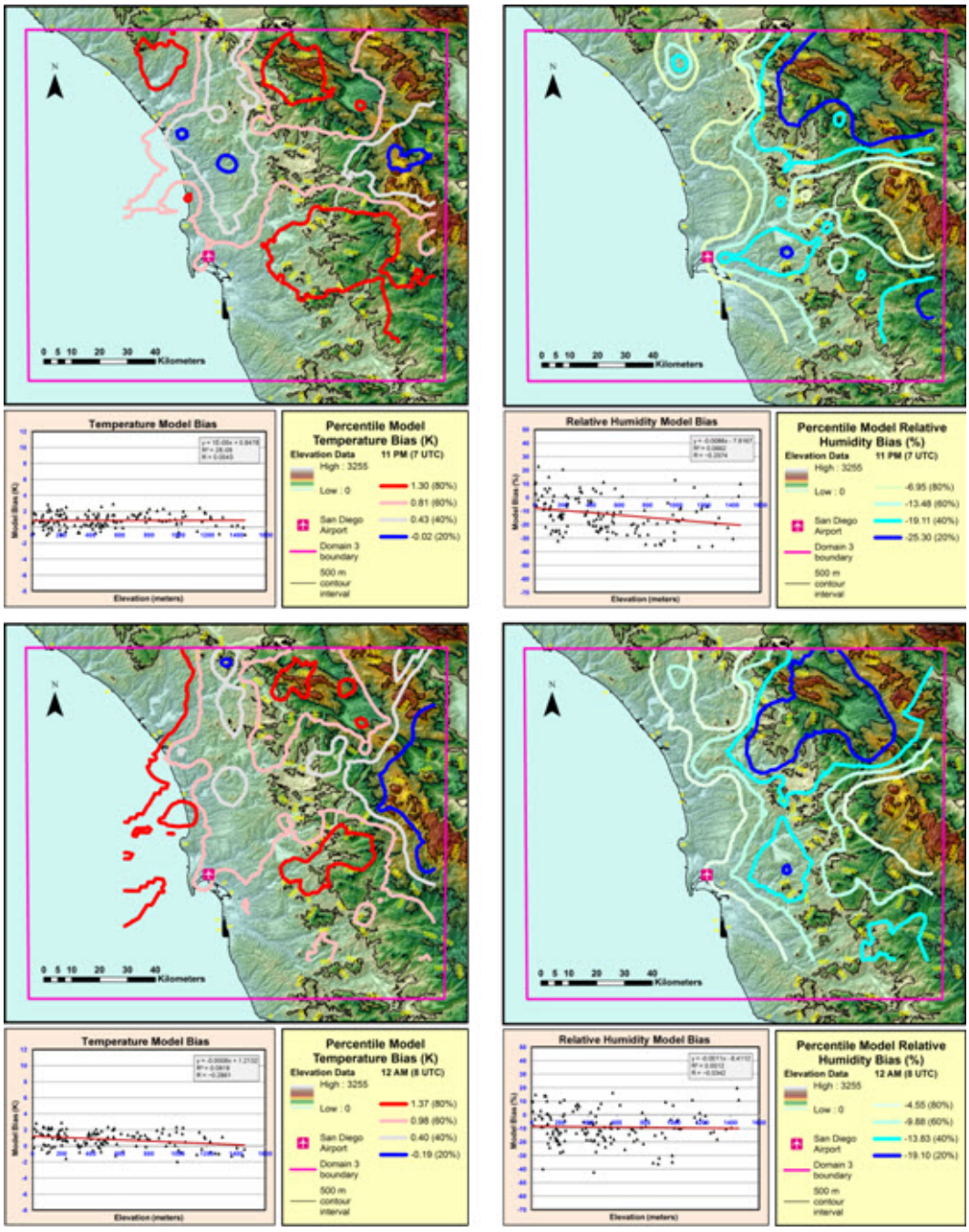

Fig. 7f EBK-created isolines of constant model error (forecast minus observed) for February 7, 2012 (case day 1) for 11:00 PM PST (7 UTC) and 12:00 AM PST (8 UTC)

Note: The temperature model error is shown in the left column and the relative humidity model error in the right column. Each map shows USGS terrain height (shaded) overlaid with colored isolines showing the quintiles of the error distribution. The scatterplot with each map shows the relationship between terrain height and model error along with a best fit line. The top row is 11:00 PM PST (7 UTC, 13-h forecast) and the bottom row is 12:00 AM PST (8 UTC, 14-h forecast).

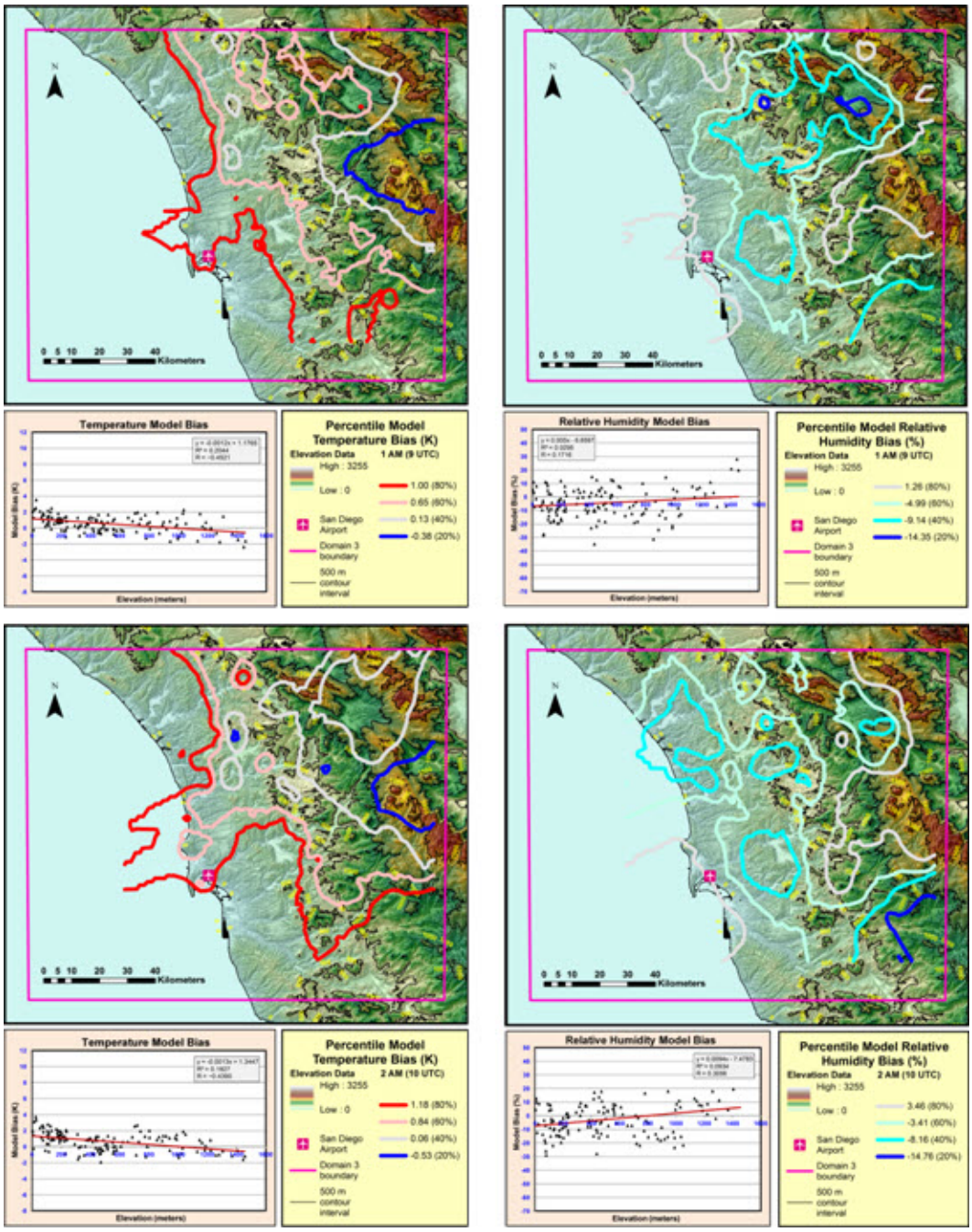

Fig. 7g EBK-created isolines of constant model error (forecast minus observed) for February 7, 2012 (case day 1) for 1:00 AM PST (9 UTC) and 2:00 AM PST (10 UTC)

Note: The temperature model error is shown in the left column and the relative humidity model error in the right column. Each map shows USGS terrain height (shaded) overlaid with colored isolines showing the quintiles of the error distribution. The scatterplot with each map shows the relationship between terrain height and model error along with a best fit line. The top row is 1:00 AM PST (9 UTC, 15-h forecast) and the bottom row is 2:00 AM PST (10 UTC, 16-h forecast).

5. Summary and Conclusions

We have found that elevation accounts for a portion of the variance in the temperature and relative humidity WRF model errors. On average, elevation accounts for about 10% to 20% of the total variance, with elevation explaining more than 50% of the total variance during the afternoon hours of February 7, 2012. By contrast, elevation accounts for none of the variance during the early morning hours of March 1 and March 5. Overall, elevation accounts for more variance during the afternoon hours and less variance during the early morning hours; but we cannot distinguish a clear diurnal pattern as to how much elevation affects the WRF model error.

Future studies will evaluate the effect of other explanatory terrain variables such as slope, aspect, and distance from the coastline, latitude, longitude, soil moisture, and land use on bias errors. Further research is needed to evaluate why terrain variables such as elevation affect the WRF model error. It would also be beneficial to evaluate how synoptic weather conditions influence the relationship between model error and terrain variables. Analysis of weather conditions would likely help us to better understand the day-to-day (and hour-to-hour) variation in the model bias errors.

Another possible source of model error may be geolocation errors. The WRF model assumes that the earth is a perfect sphere and does not use a datum to correct for the fact that the earth is not perfectly spherical. Geolocation discrepancies cause errors in the Coriolis forcing and lateral boundary conditions as well as in the locations of various observations, which can have discrepancies of as much as 20 km depending on latitude; a condition of concern for local area models with horizontal resolutions on the order of 1 km (Monaghan et al. 2013; see especially Fig. 1, p. 2121).

We achieved our goal for this study, which was to develop analytical methods that go beyond the types of conclusions gained from domain-level, aggregate statistics. Using the tools of the GIS, the model errors can be analyzed to show their possible dependence on terrain characteristics at the spatial scales of phenomena of interest to Army Warfighters—such as mountain/valley breezes and land/sea breezes. This study demonstrates that GIS tools show considerable promise toward our goal of demonstrating the relationships and dependencies of high-resolution model errors to terrain characteristics in subdomains where the statistics are more homogeneous.

6. References

Agnew MD, Palutikof JP. GIS-based construction of baseline climatologies for the mediterranean using terrain variables. Climate Research. 2000;14(2):115–127.

AirDat. Airdat Real-Time TAMDAR Weather Data and Products [Internet]. 2014. Available from: http://www.airdat.com/.

Brandt J, Dawson L, Johnson JO, Kirby S, Marlin D, Sauter D, Shirkey RC, Swanson J, Szymber R, Zeng S. Second generation weather impacts decision aid applications and web services overview. White Sands Missile Range (NM): Army Research Laboratory (US); 2013 Sep. Final Report No. ARL-TR-6525.

Brown DP, Comrie AC. Spatial modeling of winter temperature and precipitation in Arizona and New Mexico, USA. Climate Research. 2000;22(2):115–128.

Casati B, Wilson LJ, Stephenson DB, Nurmi P, Ghelli A, Pocernich M, Damrath U, Ebert EE, Brown BG, Mason S. Forecast verification: current status and future directions. Meteorological Applications. 2008;15(1):3–18.

Chapman L, Thornes JE. The use of geographical information systems in climatology and meteorology. Prog Phys Geo. 2003;27(3):313–330.

Chen F, Dudhia J. Coupling an advanced land surface–hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity. Monthly Weather Review. 2001a;129(4):569–585.

Chen F, Dudhia J. Coupling an advanced land surface–hydrology model with the Penn State-NCAR MM5 modeling system. Part II: Preliminary model validation. Monthly Weather Review. 2001b;129(4):587–604.

Cheng F–Y, Hsu Y–C, Lin P–L, Lin T–H. Investigation of the effects of different land use and land cover patterns on mesoscale meteorological simulations in the Taiwan Area. Journal of Applied Meteorology and Climatology. 2012;52(3):570–587.

Daniels TS, Murray JJ, Grainger CA, Zhou DK, Avery MA, Cagle MF, Tsoucalas G, Schaffner PR, Neece RT. 2004. Validation of tropospheric airborne meteorological data reporting (TAMDAR) temperature, relative humidity, and wind sensors during the 2003 Atlantic Thorpex Regional Campaign and the Alliance Icing Research Study (AIRS II). 11th AMS Conference on Aviation, Range, and Aerospace. Hyannis (MA): American Meteorological Society.

Deng A, Stauffer D, Gaudet B, Dudhia J, Bruyere C, Wu W, Vandenberghe F, Liu Y, Bourgeoi A. 2009. Update on the WRF–ARW end-to-end multiscale FDDQA system. 10th WRF Users' Workshop. Boulder, CO: National Center for Atmospheric Research. p. paper 1.9.

Dudhia J. Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. Journal of the Atmospheric Sciences 1989;46(20):3077–3107.

Dumais R, Kirby S, Flanigan R. Implementation of the WRF four-dmensional data assimilation method of observation nudging for use as an ARL weather running estimate-nowcast. White Sands Missile Range (NM): Army Research Laboratory (US); 2013 Jun. Final Report No. ARL-TR-6485. Also available at

http://www.arl.army.mil/www/default.cfm?technical_report=6767.

Dumais RE, Jr., Henmi T, Passner J, Jameson T, Haines P, Knapp D. A mesoscale modeling system developed for the US Army. White Sands Missile Range (NM): Army Research Laboratory (US); 2004 Apr. Final Report No. ARL-TR-3183.

Dumais RE, Raby JW, Wang Y, Raby YR, Knapp D. Performance assessment of the three-dimensional wind field weather running estimate-nowcast and the three-dimensional wind field Air Force Weather Agency weather research and forecasting wind forecasts. White Sands Missile Range (NM): Army Research Laboratory (US); 2012 Dec. Technical Note No. ARL-TN-0514. Also available at

http://www.arl.army.mil/www/default.cfm?technical_report=6648.

Dumais RE, Reen BP. Data assimilation techniques for rapidly relocatable weather research and forecasting modeling. White Sands Missile Range (NM): Army Research Laboratory (US); 2013. Final Report No. ARL-TN-0546.

Ebert E, Wilson L, Weigel A, Mittermaier M, Nurmi P, Gill P, Gober M, Joslyn S, Brown B, Fowler T et al. Progress and challenges in forecast verification. Meteorological Applications. 2013;20(2):130–139.

Environmental Modeling Center 2003. The GFS atmospheric model. Washington, DC: National Oceanic and Atmospheric Administration, National Weather Service, National Centers for Environmental Prediction. National Centers for Environmental Prediction Office Note No. 442.

ESRI. 2014. ArcGIS for desktop with ArcGIS geostatistical analysis, spatial statistics and 3D analysis toolboxes. ArcGIS 10.2 ed.: Environmental Systems Research Institute, Inc. p. ArcGIS for Desktop with ArcGIS Geostatistical Analysis, Spatial Statistics and 3D Analysis Toolboxes.

Hong S–Y, Dudhia J, Chen S–H. A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. Monthly Weather Review. 2004;132(1):103–120.

Janjić ZI. The step-mountain ETA coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. Monthly Weather Review. 1994;122(5):927–945.

Johnson JO. Statistical analysis of atmospheric forecast model accuracy for multiple levels of forecast model resolution – A focus on multiple atmospheric variables and location-based analysis. Las Cruces, NM: New Mexico State University, 2012.

Johnson JO, Raby JW, Knapp DI. Statistical analysis of atmospheric forecast model accuracy – A focus on multiple atmospheric variables and location-based analysis. White Sands Missile Range (NM): Army Research Laboratory (US); 2014. Final Report No. ARL-TR-6915.

Jolliffe IT, Stephenson DB. Forecast verification: A practitioner's guide in atmospheric science. Hoboken (NJ): John Wiley & Sons; 2011.

Kain JS. The Kain-Fritsch Convective Parameterization: An Update. Journal of Applied Meteorology. 2004;43(1):170–181.

Krivoruchko K. Spatial Statistical Data Anlysis for GIS Users. Redlands (CA): ESRI Press; 2011.

Krivoruchko K. Empirical Bayesian Kriging: Implemented in ArcGIS Geostatistical Analyst. ArcUser. 2012;15(4):6–10.

Liu Y, Bourgeois A, Warner T, Swerdlin S, Hacker J. Implementation of observation-nudging based FDDA into WRF for supporting ATEC test operations. The 6th WRF/15th MM5 Users Workshop. Boulder, CO: National Center for Atmospheric Research; 2005.

Massey JD, Steenburgh WJ, Hoch SW, Knievel JC. 2014. Sensitivity of near-surface temperature forecasts to soil properties over a sparsely vegetated dryland region. Journal of Applied Meteorology and Climatology. 2014;53(8):1976–1995.

Mlawer EJ, Taubman SJ, Brown PD, Iacono MJ, Clough SA. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-K model for the longwave. Journal of Geophysical Research-Atmospheres. 1997;102(D14):16663–16682.

Monaghan AJ, Barlage M, Boehnert J, Phillips CL, Wilhelmi OV. Overlapping interests: The impact of geographic coordinate assumptions on limited-area atmospheric model simulations. Monthly Weather Review. 2013;141(6):2120–2127.

National Center for Atmospheric Research. Model Evaluation Tools Version 4.1 (Metv4.1). Boulder, CO. User's Guide No. 4.1; 2013.

National Research Council. 2010. When Weather Matters: Science and Service to Meet Critical Societal Needs. Washington (DC): The National Academies Press.

NOAA. Meteorological Assimilation Data Ingest System (MADIS) [Internet]. [cited 2014. Available from: http://madis.noaa.gov/.

NOAA. Real-Time Mesoscale Analysis (RTMA) [Internet]. [cited 2014. Available from: http://www.nco.ncep.noaa.gov/pmb/products/rtma/.

Raby J, Passner J, Vaucher G, Raby Y. Performance comparison of high resolution weather research and forecasting model output with North American mesoscale model initialization grid forecasts. White Sands Missile Range (NM): Army Research Laboratory (US); 2012 May. Final Report No. ARL-TR-6000.

Reen BP, Schmehl KJ, Young GS, Lee JA, Haupt SE, Stauffer DR. Uncertainty in contaminant concentration fields resulting from atmospheric boundary layer depth uncertainty. Journal of Applied Meteorology and Climatology. 2014;53(11):2610–2626.

Ross J. The forecast for D-day and the weatherman behind Ike's greatest gamble. Guilford (CT): Lyons Press, an imprint of Globe Pequot Press; 2014.

Skamarock WC, Klemp JB, Dudhia J, Gill DO, Barker DM, Duda MG, Huang X–Y, Wang W, Powers JG. A description of the advanced research WRF version 3. Boulder, CO. NCAR Technical Note No. NCAR/TN-475+STR; 2008.

Smith JA, Foley T, Raby J. Using a geographic information system (GIS) to validate high resolution weather forecasts. Fall Meeting of the American Geophysical Union; 2014a; San Francisco, CA. p. Poster A43F-3350; 2014.

Smith JA, Foley TA, Raby JW. Developing subdomain verification methods based on GIS tools. 15th WRF Users Workshop; 2014b; Boulder, CO; 2014.

Smith JA, Foley TA, Raby JW. Developing subdomain verification methods based on GIS tools. 95th Annual AMS Meeting, Special Symposium Model Post Processing and Downscaling; 2015; Phoenix, AZ. p. Poster 846; 2015.

Smith JA, Foley TA, Raby JW, Reen B. Developing subdomain verification methods based on GIS tools. 16th Mountain Meteorology Conference; 2014c; San Diego, CA; 2014.

Smith JA, Raby J, Foley T. Comparing high resolution weather forecasts to observations. Fall Meeting of the American Geophysical Union; 2013.

Stauffer DR, Seaman NL. Multiscale four-dimensional data assimilation. Journal of Applied Meteorology 1994;33(3):416–434.

US Geological Survey. Digital Elevation Model. 2013.

Wilhelmi O, Boehnert J. 2014. Personal Communication. In: Foley T, editor.

Wilks DS. 2011. Statistical methods in the atmospheric sciences. Oxford: Academic Press.

Appendix A. Domain Level Errors for All 5 Case Days

Table A-1 Mean bias (modeled−observed) errors for each of the 5 case days, and the 3 meteorological variables modeled at the Z2 surface level (2 meters AGL). These statistics were calculated over the entire domain.

February 7–8, 2012				
Hour (PST)	Hour (UTC)	Temperature (K)	Dew Point Temp. (K)	Relative Humidity (%)
4:00 AM	12	0.09	1.06	0.96
5:00 AM	13	−0.08	−0.16	−1.37
10:00 AM	18	0.69	−0.94	−5.70
4:00 PM	0	−0.17	1.11	5.83
10:00 PM	6	0.35	−2.22	−12.17
4:00 AM	12	0.76	1.40	2.39
February 9–10, 2012				
Hour (PST)	Hour (UTC)	Temperature (K)	Dew Point Temp. (K)	Relative Humidity (%)
4:00 AM	12	0.17	0.17	−0.38
5:00 AM	13	−0.57	0.82	3.49
10:00 AM	18	−0.06	0.96	2.56
4:00 PM	0	−2.89	2.95	12.27
10:00 PM	6	1.33	−0.18	−5.28
4:00 AM	12	2.42	−1.35	−13.76
February 16–17, 2012				
Hour (PST)	Hour (UTC)	Temperature (K)	Dew Point Temp. (K)	Relative Humidity (%)
4:00 AM	12	0.14	0.06	−0.46
5:00 AM	13	0.59	0.14	−2.64
10:00 AM	18	0.85	−0.40	−5.35
4:00 PM	0	−0.65	0.09	1.94
10:00 PM	6	0.10	−1.29	−5.63
4:00 AM	12	0.81	−2.18	−11.31
March 1–2, 2012				
Hour (PST)	Hour (UTC)	Temperature (K)	Dew Point Temp. (K)	Relative Humidity (%)
4:00 AM	12	0.15	−0.03	−0.40
5:00 AM	13	−0.02	0.19	1.84
10:00 AM	18	0.41	0.25	0.07
4:00 PM	0	0.30	0.23	0.14
10:00 PM	6	−0.54	−0.87	−1.24
4:00 AM	12	1.58	−2.18	−18.66
March 5–6, 2012				
Hour (PST)	Hour (UTC)	Temperature (K)	Dew Point Temp. (K)	Relative Humidity (%)
4:00 AM	12	0.81	−0.33	−2.51
5:00 AM	13	−0.63	−0.82	−1.09
10:00 AM	18	−0.03	0.33	1.31
4:00 PM	0	−3.55	3.48	17.34
10:00 PM	6	0.99	−2.23	−10.77
4:00 AM	12	−0.13	−0.59	−1.57

INTENTIONALLY LEFT BLANK.

Appendix B. Remaining Case Days beyond Fig. 6

Note: Each map in this appendix shows US Geological Survey (USGS) terrain height (shaded) overlaid with colored isolines showing the quintiles of the bias distribution. The left panel shows a scatterplot of the relationship between terrain height and model bias along with a best fit line.

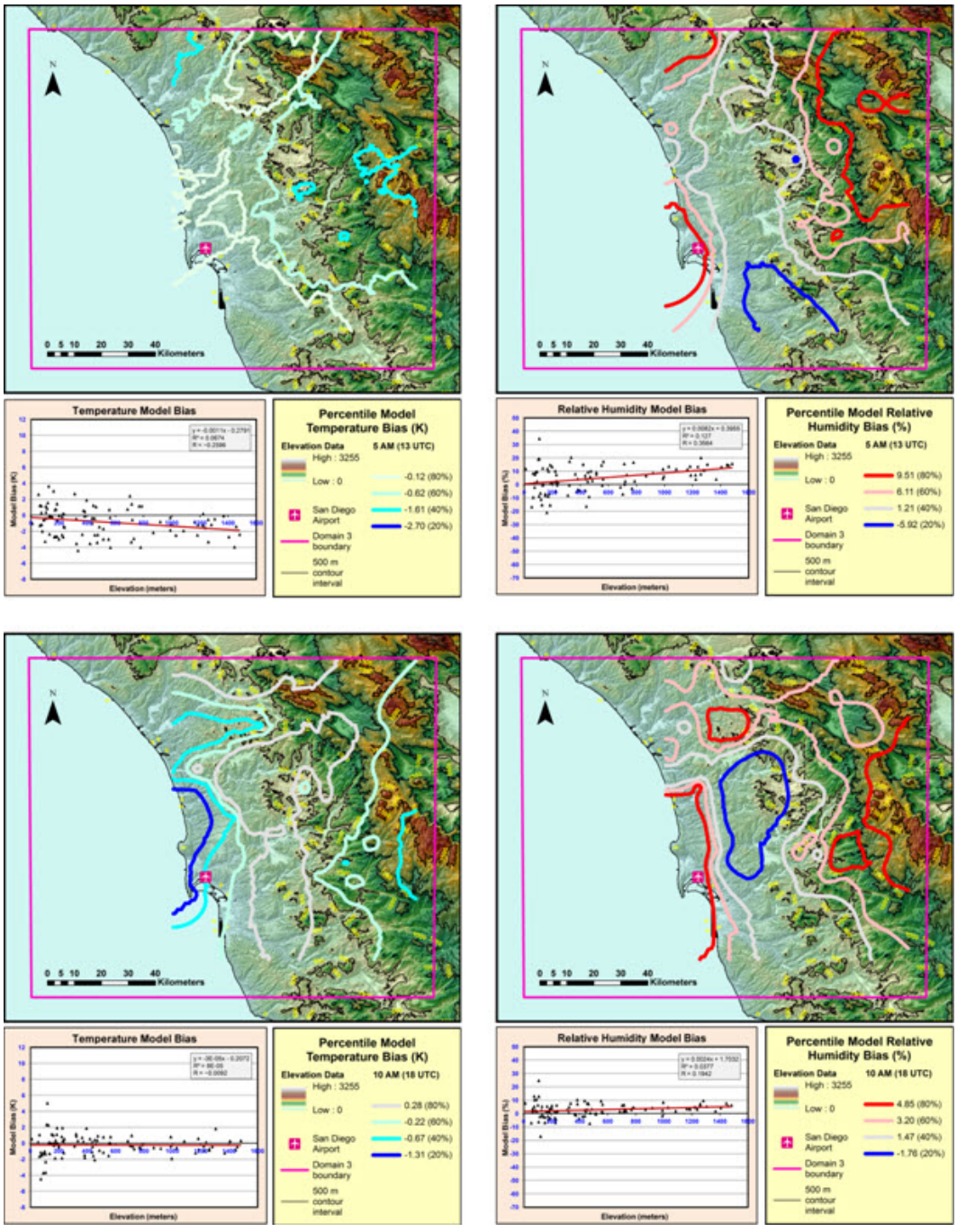

Fig. B-1a Weather Research and Forecasting (WRF) 1-km domain model surface bias (forecast minus observed) for the February 9, 2012, case for temperature (left column) and relative humidity (right column). The top row is 5:00 AM Pacific Standard Time (PST) (13 coordinated universal time [UTC], 1 h into preforecast) and the bottom row is 10:00 AM PST (18 UTC, 0-h forecast).

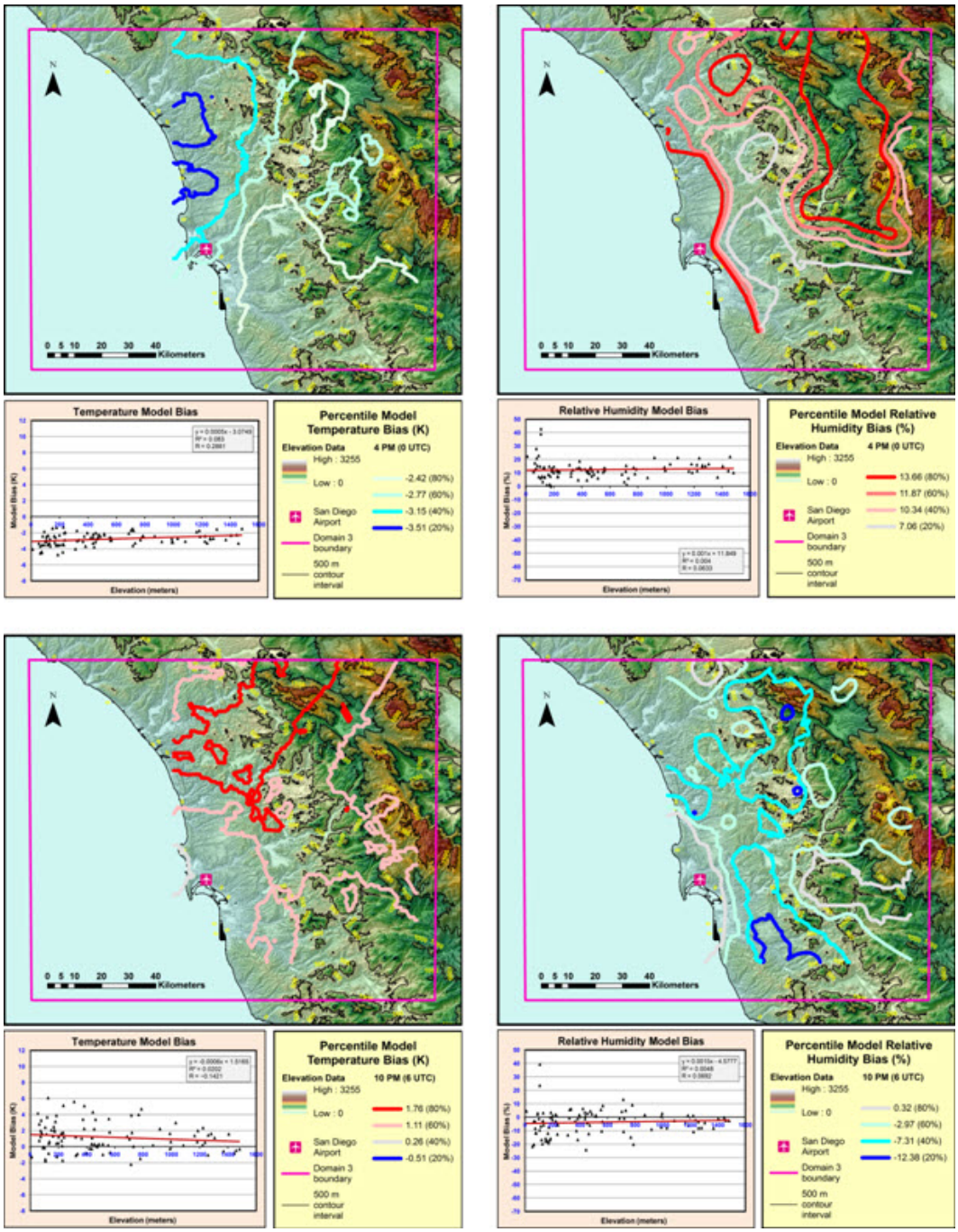

Fig. B-1b WRF 1-km domain model surface bias (forecast minus observed) for the February 9, 2012, case for temperature (left column) and relative humidity (right column). The top row is 4:00 PM PST (00 UTC, 6-h forecast) and the bottom row is 10:00 PM PST (06 UTC, 12-h forecast)

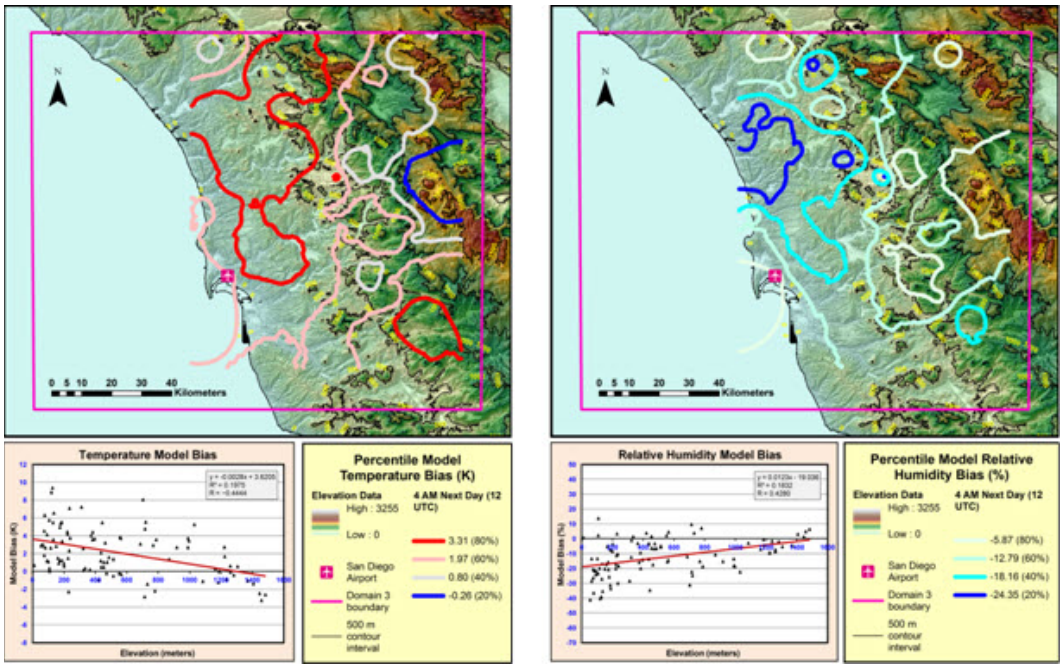

Fig. B-1c WRF 1-km domain model surface bias (forecast minus observed) for the February 9, 2012, case for temperature (left figure) and relative humidity (right figure), 4:00 AM the next day (12 UTC, 18-h forecast)

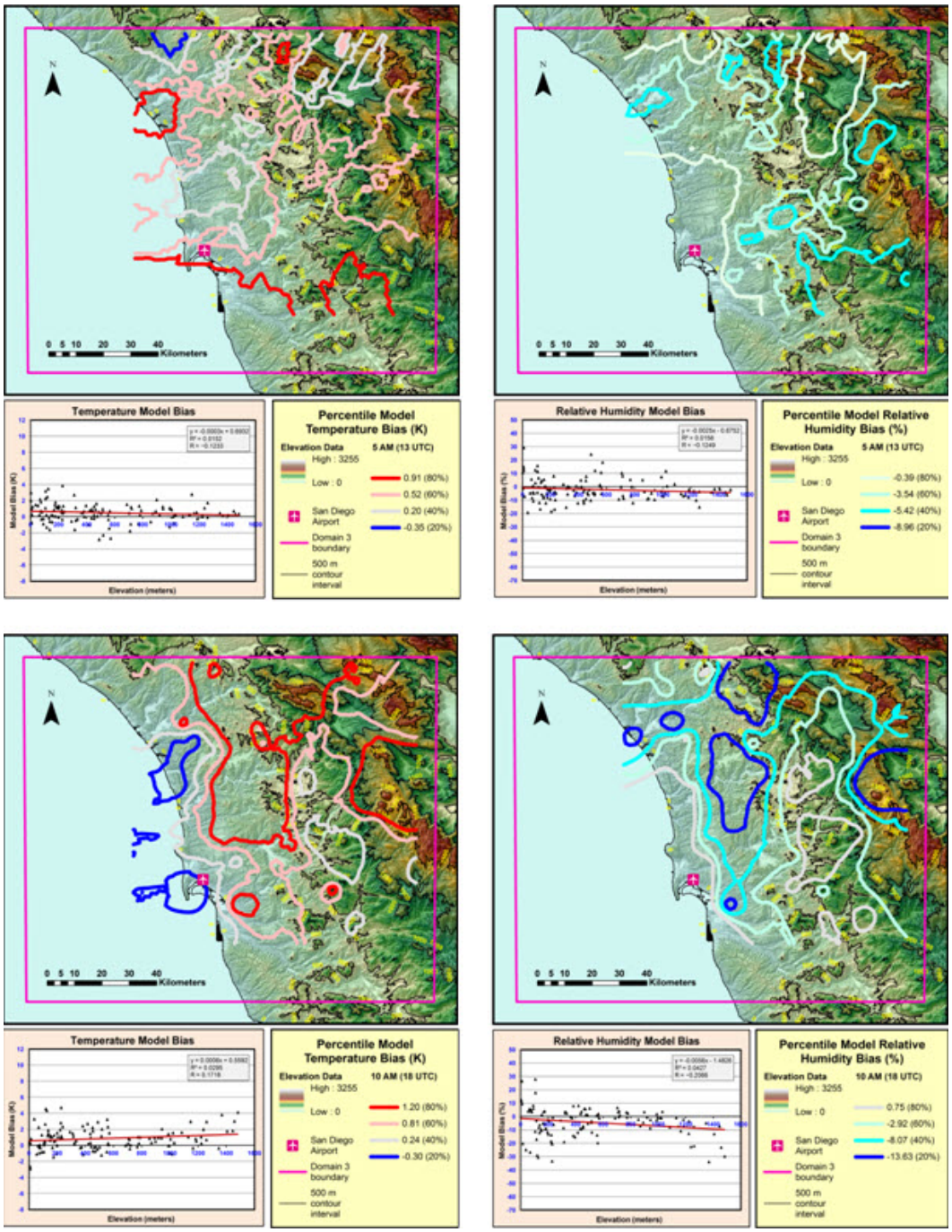

Fig. B-2a WRF 1-km domain model surface bias (forecast minus observed) for the February 16, 2012, case for temperature (left column) and relative humidity (right column). The top row is 5:00 AM PST (13 UTC, 1 h into preforecast) and the bottom row is 10:00 AM PST (18 UTC, 0–h forecast).

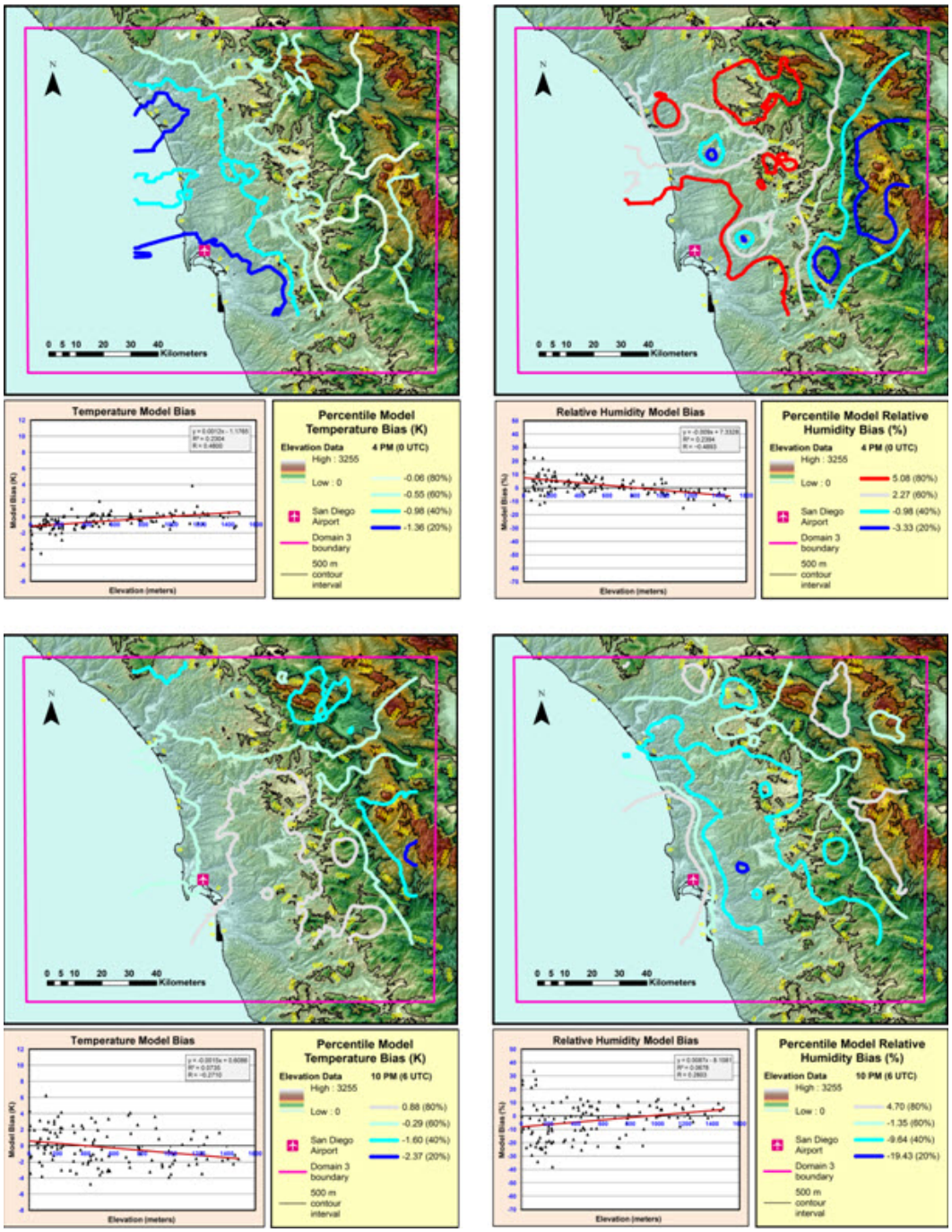

Fig. B-2b WRF 1-km domain model surface bias (forecast minus observed) for the February 16, 2012, case for temperature (left column) and relative humidity (right column). The top row is 4:00 PM PST (00 UTC, 6-h forecast) and the bottom row is 10:00 PM PST (06 UTC, 12-h forecast).

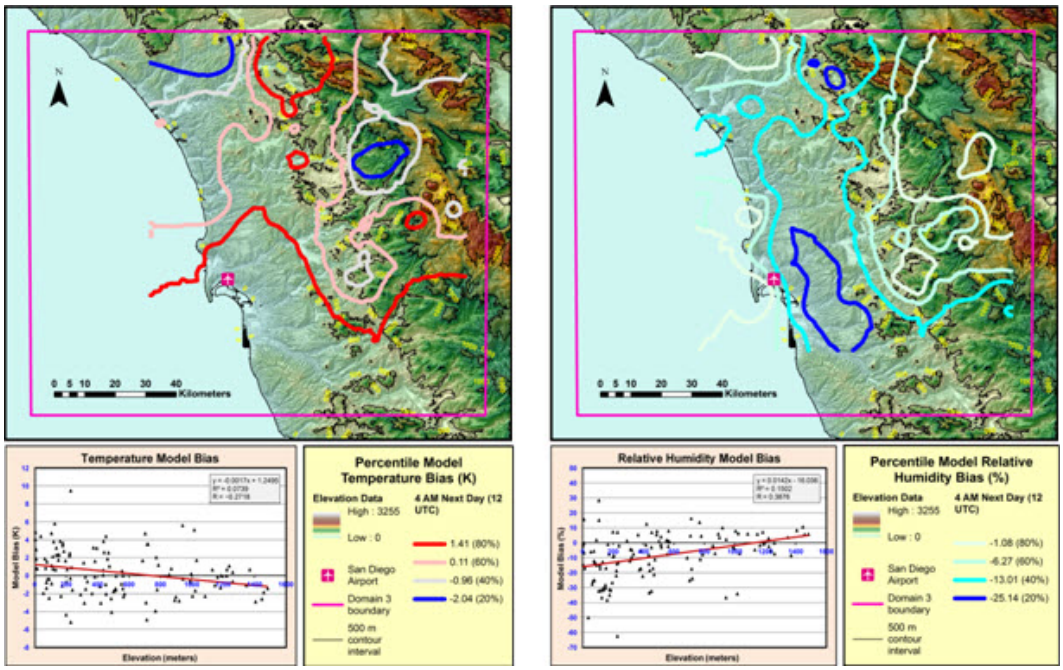

Fig. B-2c WRF 1-km domain model surface bias (forecast minus observed) for the February 16, 2012, case for temperature (left figure) and relative humidity (right figure), 4:00 AM the next day (12 UTC, 18-h forecast)

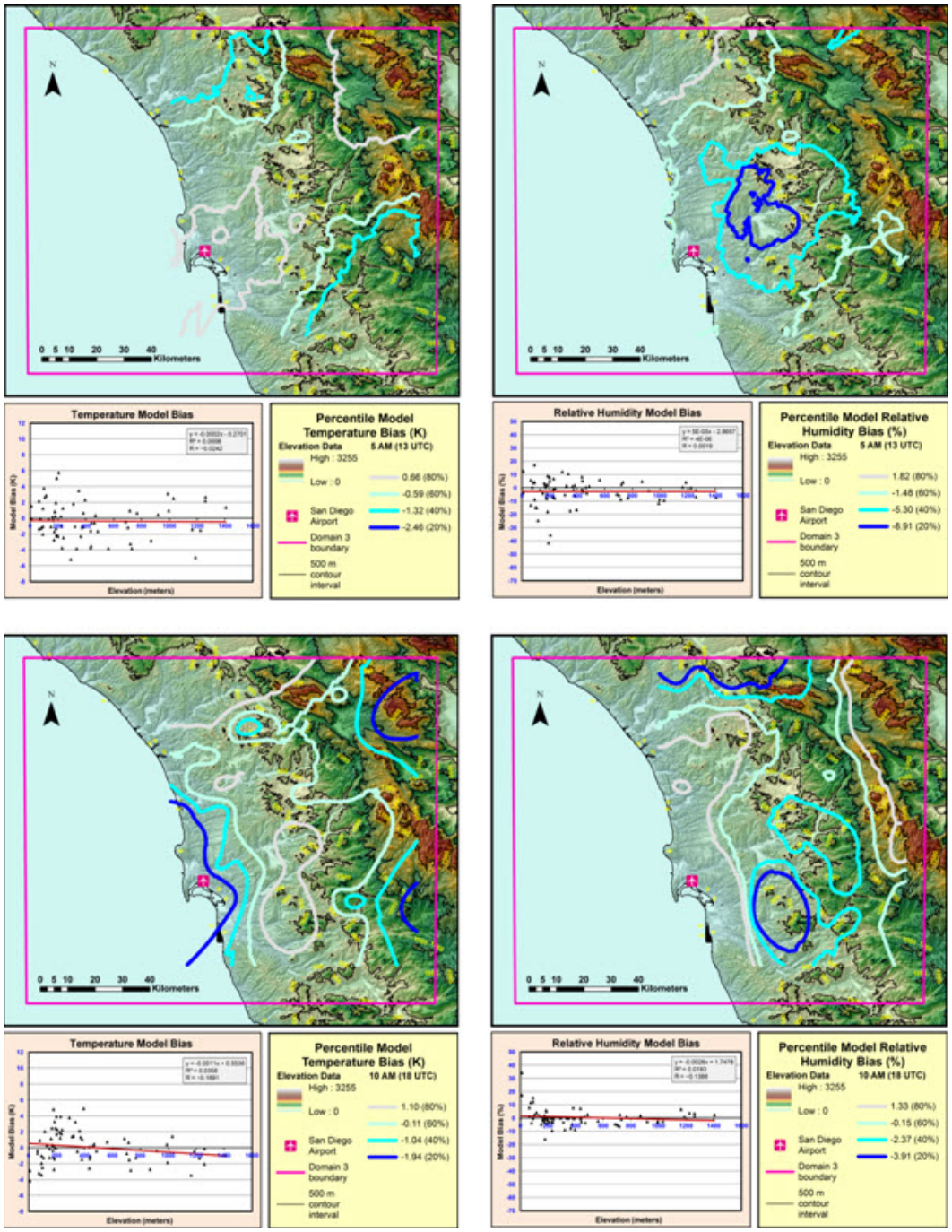

Fig. B-3a WRF 1-km domain model surface bias (forecast minus observed) for the March 5, 2012, case for temperature (left column) and relative humidity (right column). The top row is 5:00 AM PST (13 UTC, 1 h into preforecast) and the bottom row is 10:00 AM PST (18 UTC, 0–h forecast).

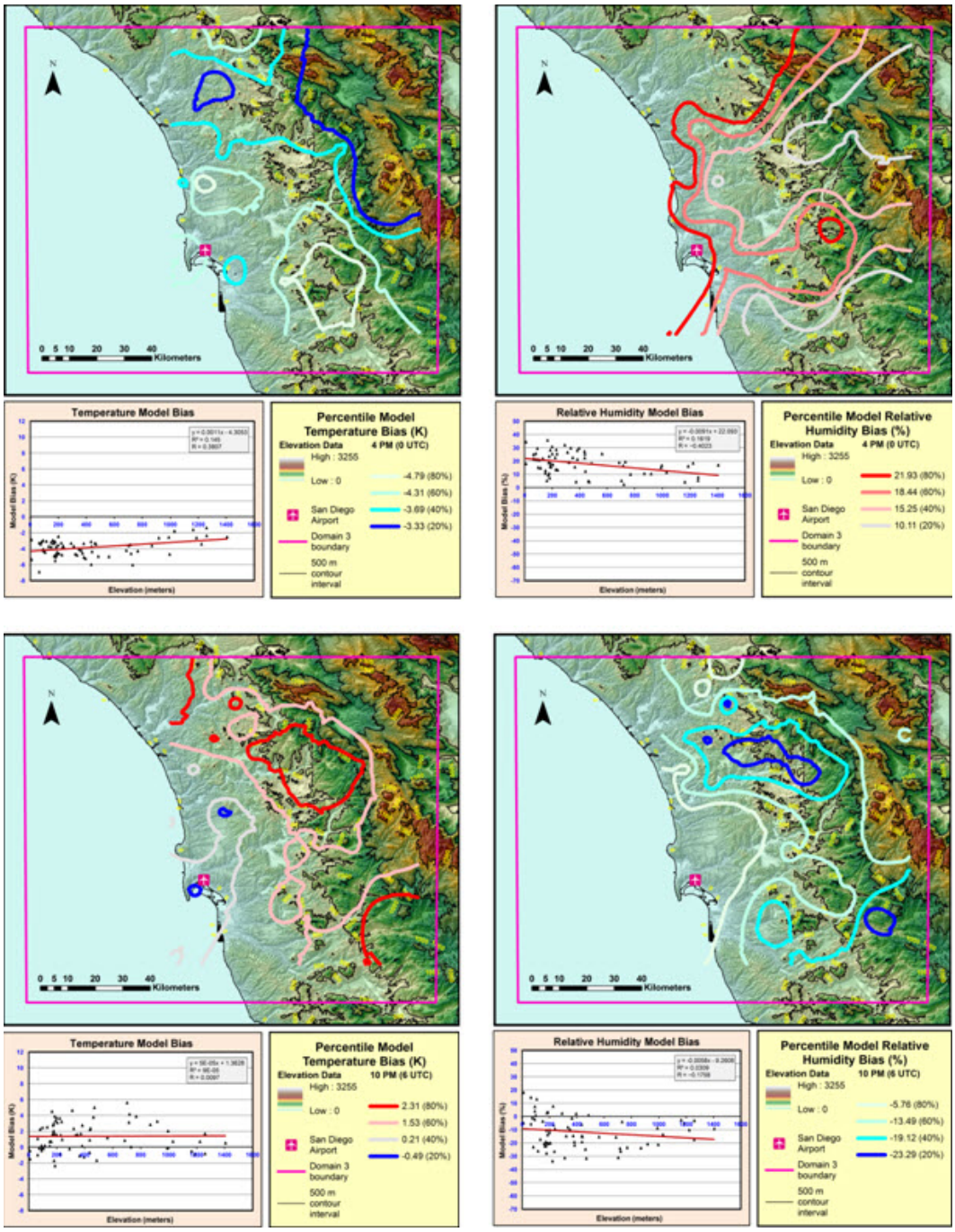

Fig. B-3b WRF 1-km domain model surface bias (forecast minus observed) for the March 5, 2012, case for temperature (left column) and relative humidity (right column). The top row is 4:00 PM PST (00 UTC, 6-h forecast) and the bottom row is 10:00 PM PST (06 UTC, 12-h forecast).

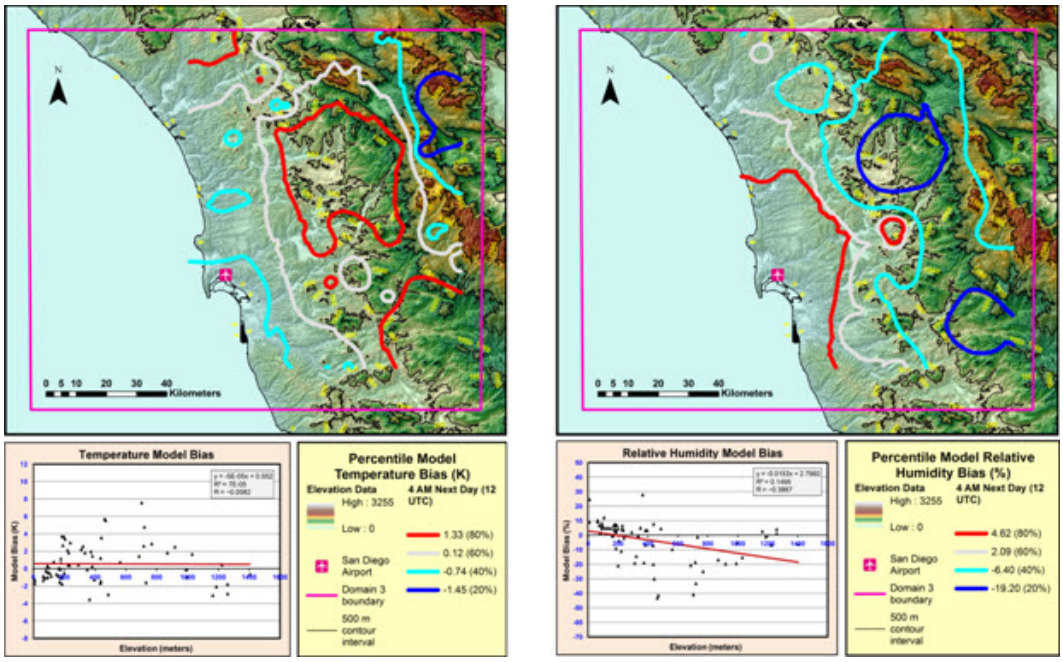

Fig. B-3c WRF 1-km domain model surface bias (forecast minus observed) for the March 5, 2012, case for temperature (left figure) and relative humidity (right figure), 4:00 AM the next day (12 UTC, 18-h forecast)

Appendix C. Case Day 1, All Hours

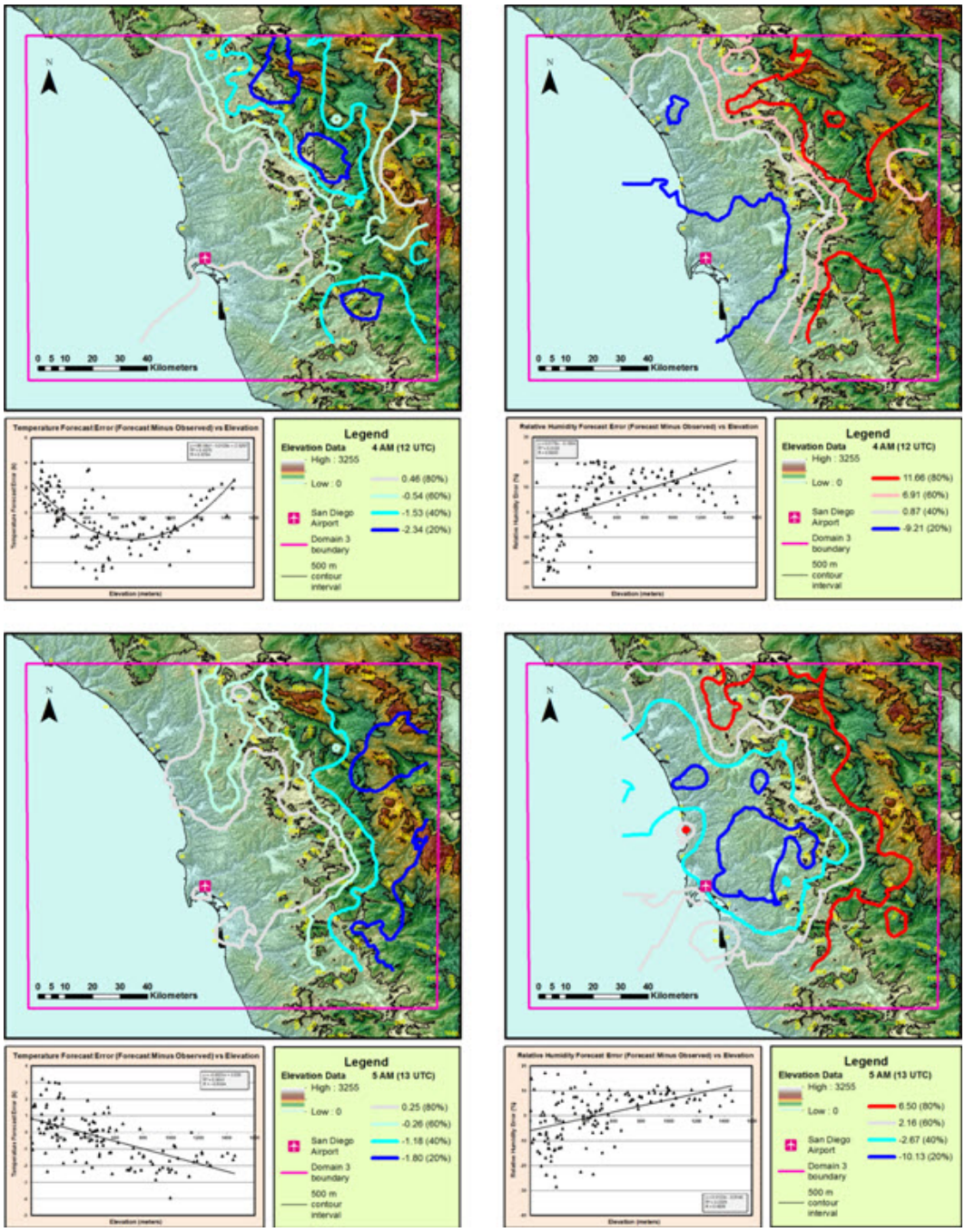

Fig. C-1 February 7, 2012, Weather Research and Forecasting (WRF) error (forecast minus observed). Elevation versus temperature forecast error (left column) and elevation versus relative humidity forecast error (right column). The top row is 4:00 AM Pacific Standard Time (PST) (12 UTC) and the bottom row is 5:00 AM PST (13 UTC).

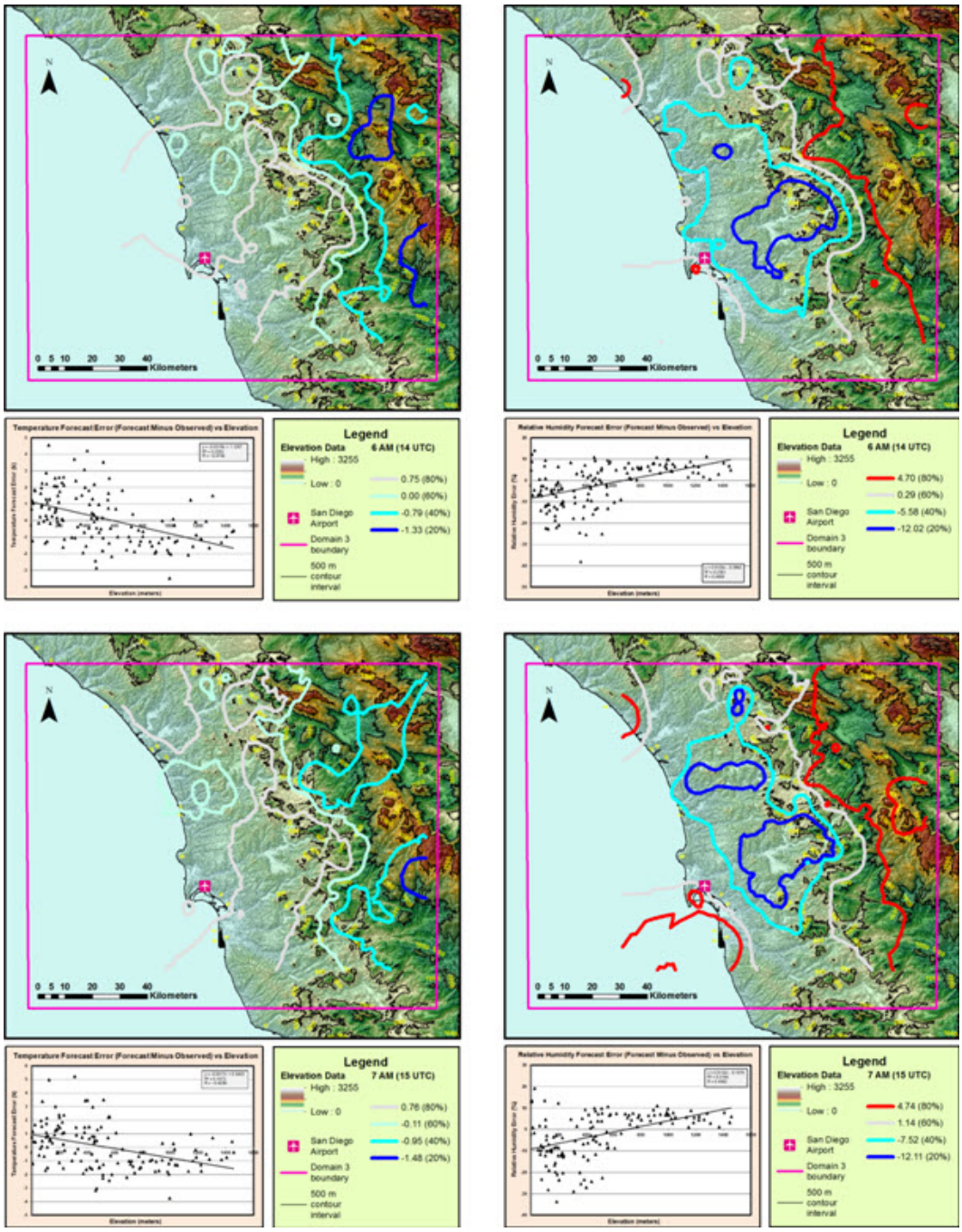

Fig. C-2 February 7, 2012, WRF error (forecast minus observed). Elevation versus temperature forecast error (left column) and elevation versus relative humidity forecast error (right column). The top row is 6:00 AM PST (14 UTC), and the bottom row is 7:00 AM PST (15 UTC).

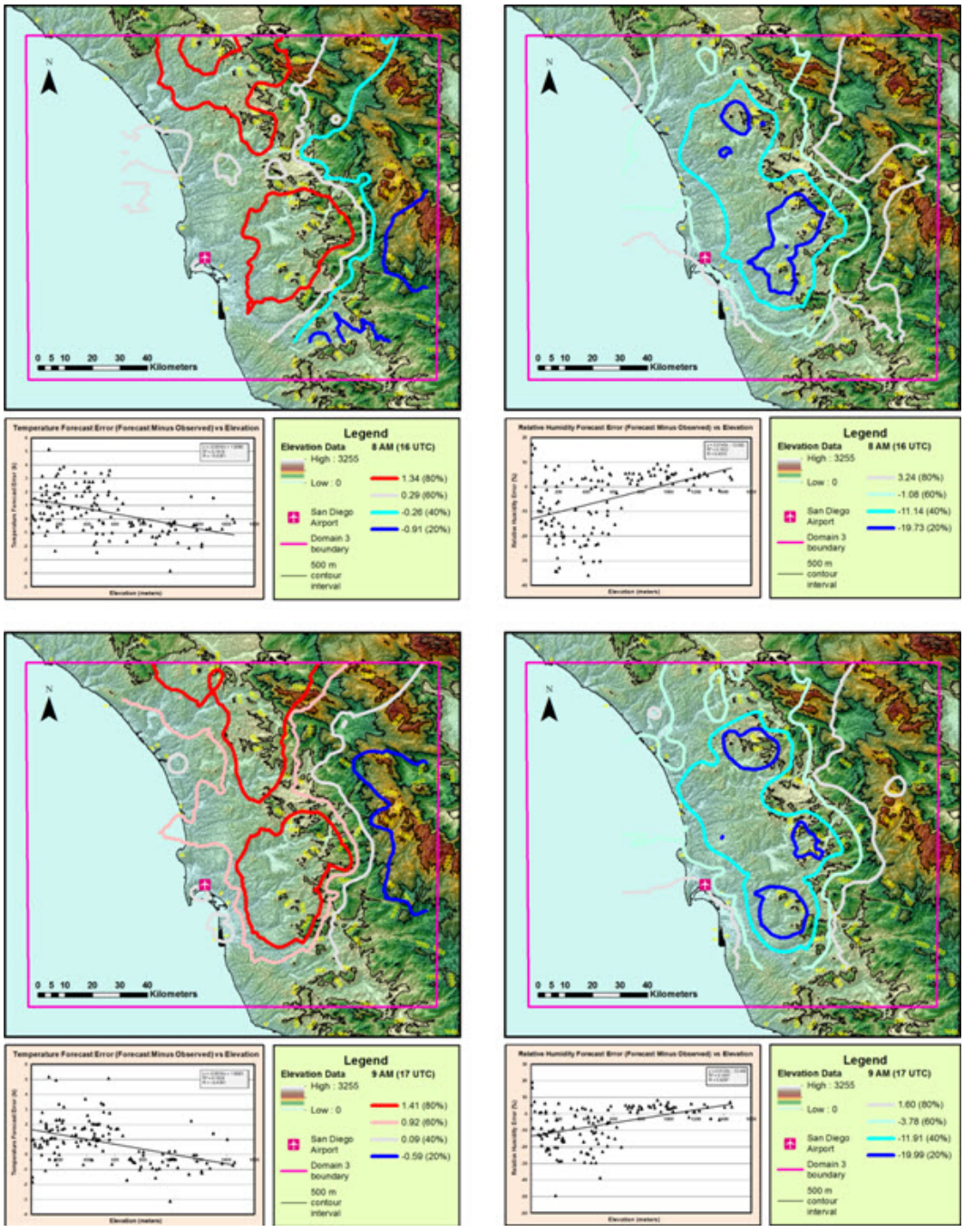

Fig. C-3 February 7, 2012, WRF error (forecast minus observed). Elevation versus temperature forecast error (left column) and elevation versus relative humidity forecast error (right column). The top row is 8:00 AM PST (16 UTC), and the bottom row is 9:00 AM PST (17 UTC).

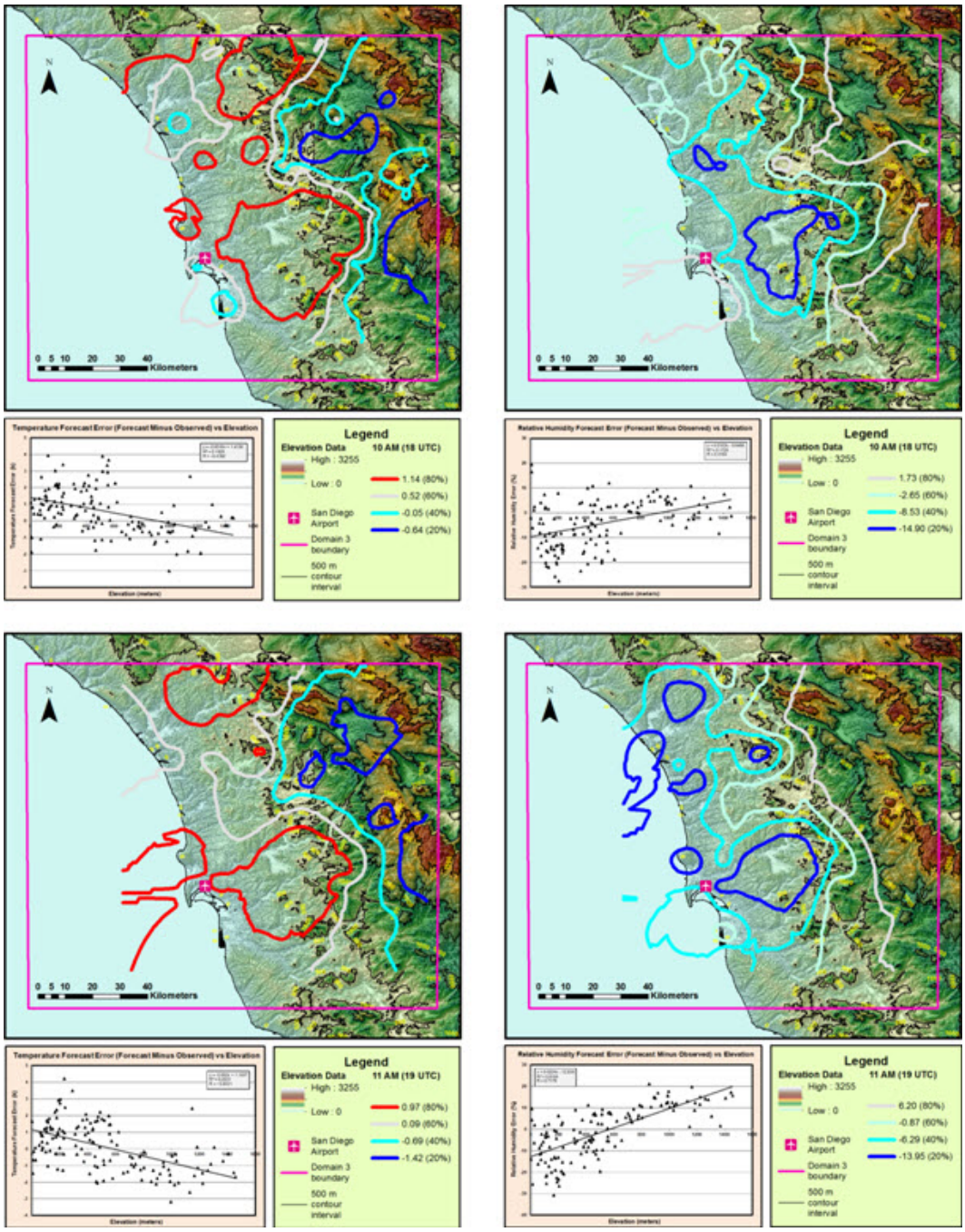

Fig. C-4 February 7, 2012, WRF error (forecast minus observed). Elevation versus temperature forecast error (left column) and elevation versus relative humidity forecast error (right column). The top row is 10:00 AM PST (18 UTC), and the bottom row is 11:00 AM PST (19 UTC).

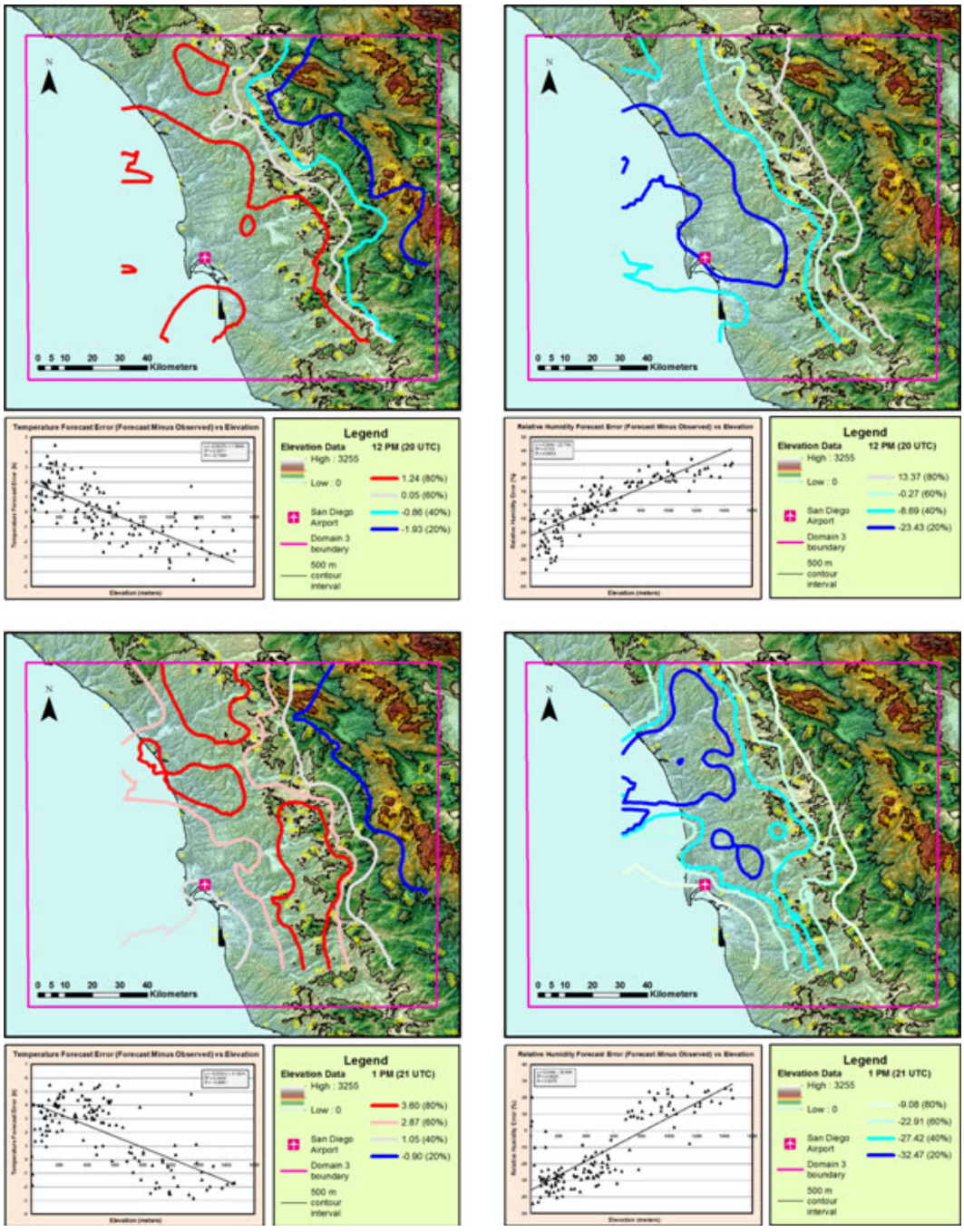

Fig. C-5 February 7, 2012, WRF error (forecast minus observed). Elevation versus temperature forecast error (left column) and elevation versus relative humidity forecast error (right column). The top row is 12:00 PM PST (20 UTC), and the bottom row is 1:00 PM PST (21 UTC).

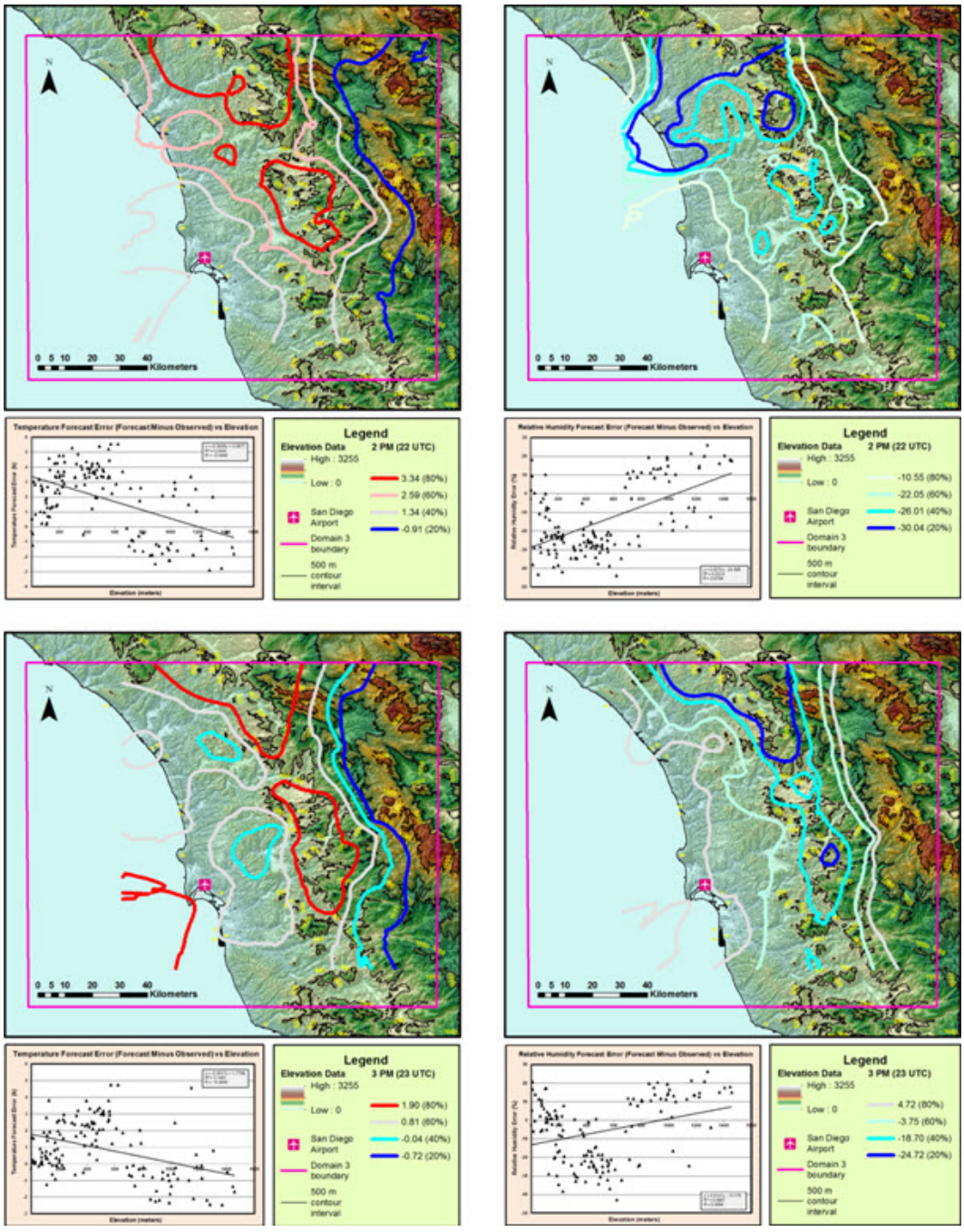

Fig. C-6 February 7, 2012, WRF error (forecast minus observed). Elevation versus temperature forecast error (left column) and elevation versus relative humidity forecast error (right column). The top row is 2:00 PM PST (22 UTC), and the bottom row is 3:00 PM PST (23 UTC).

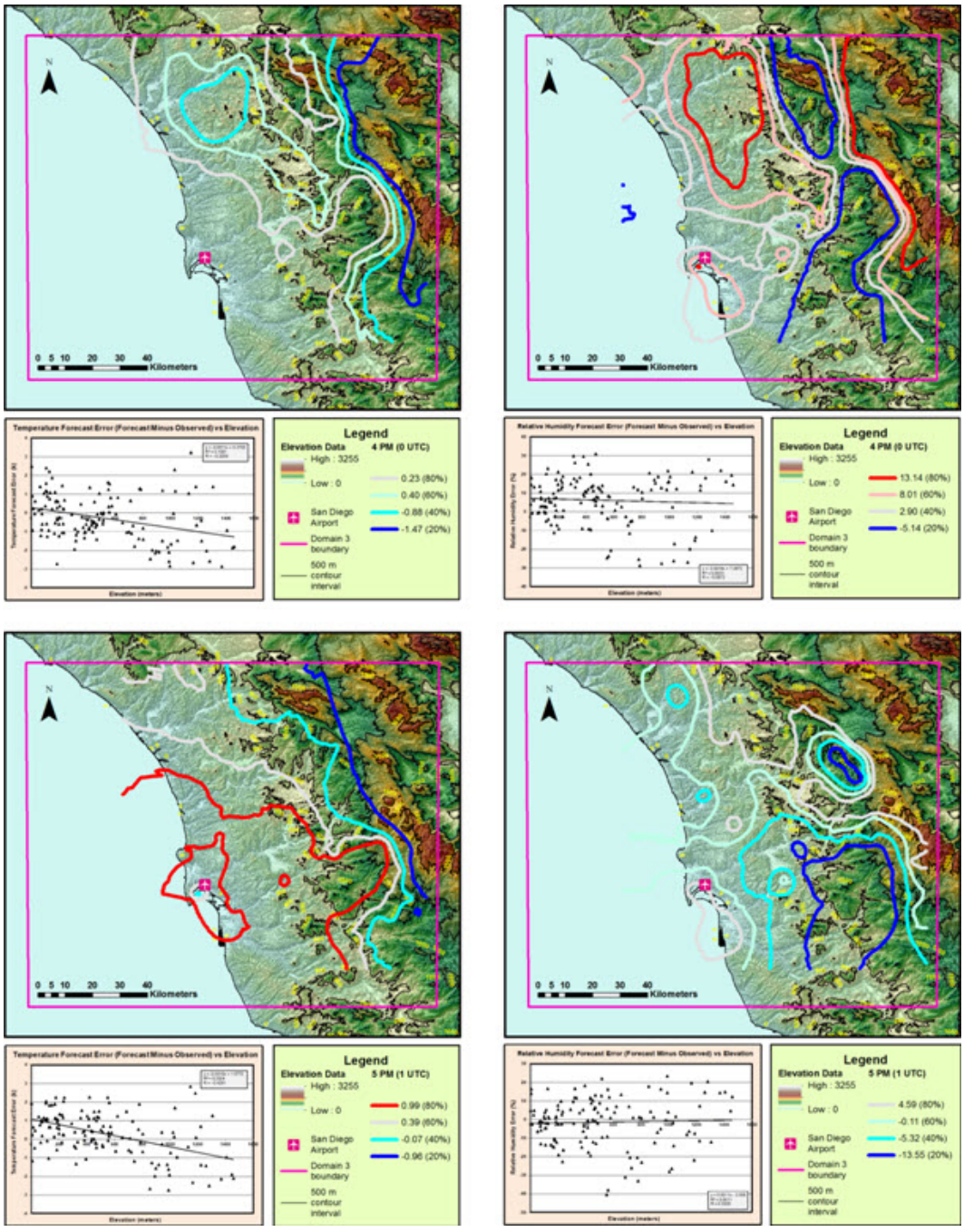

Fig. C-7 February 7, 2012, WRF error (forecast minus observed). Elevation versus temperature forecast error (left column) and elevation versus relative humidity forecast error (right column). The top row is 4:00 PM PST (0 UTC), and the bottom row is 5:00 PM PST (1 UTC).

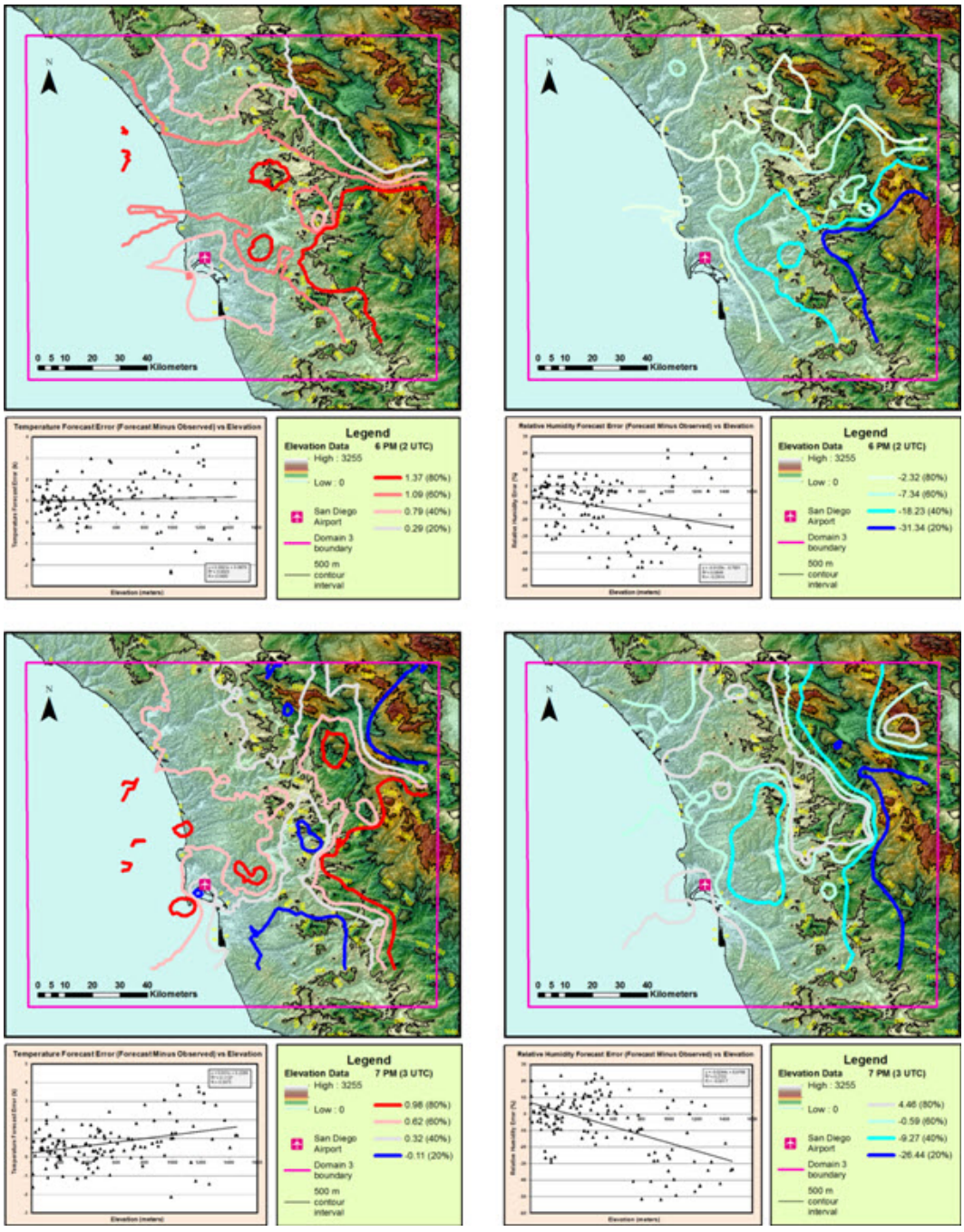

Fig. C-8 February 7, 2012, WRF error (forecast minus observed). Elevation versus temperature forecast error (left column) and elevation versus relative humidity forecast error (right column). The top row is 6:00 PM PST (2 UTC), and the bottom row is 7:00 PM PST (3 UTC).

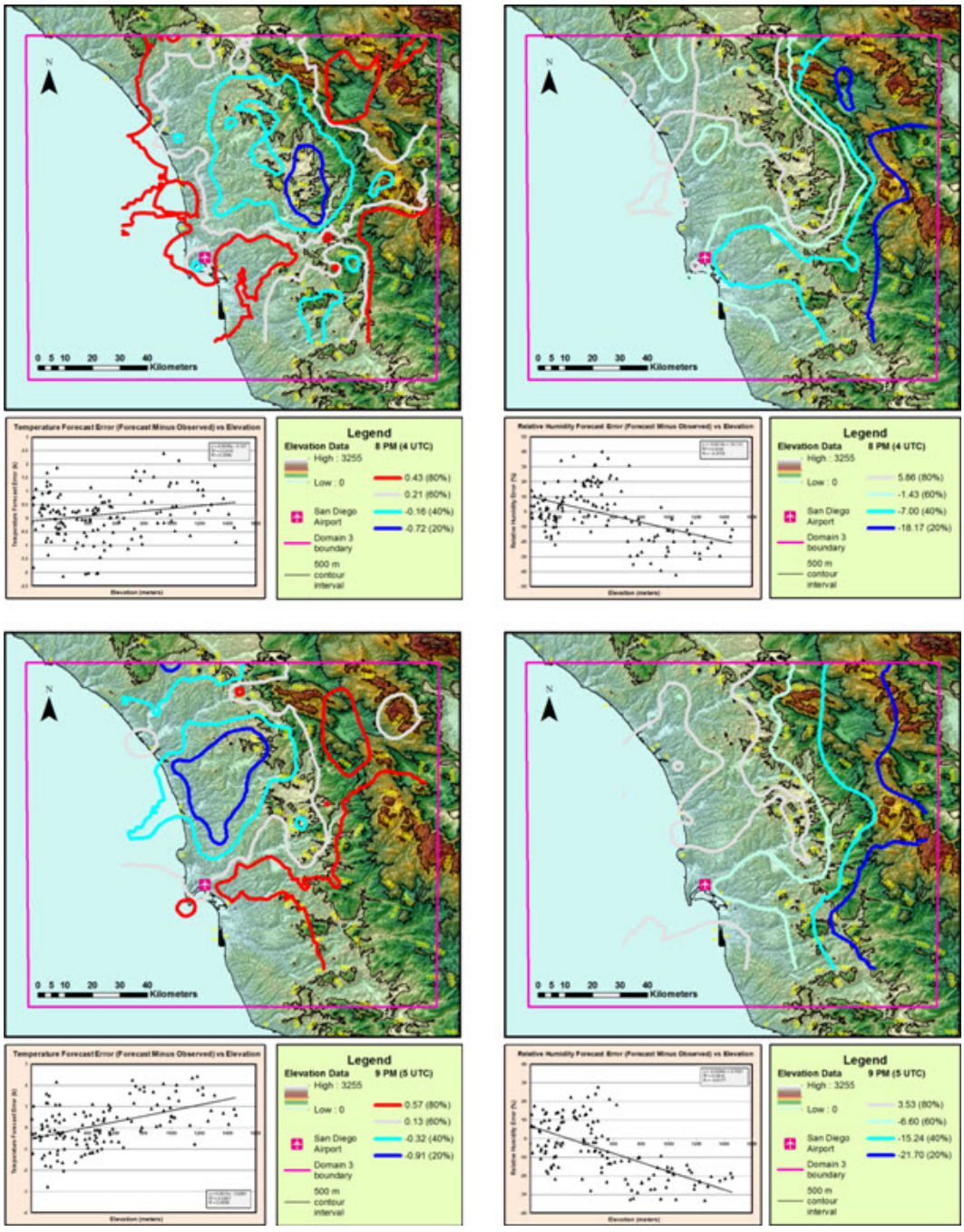

Fig. C-9 February 7, 2012, WRF error (forecast minus observed). Elevation versus temperature forecast error (left column) and elevation versus relative humidity forecast error (right column). The top row is 8:00 PM PST (4 UTC), and the bottom row is 9:00 PM PST (5 UTC).

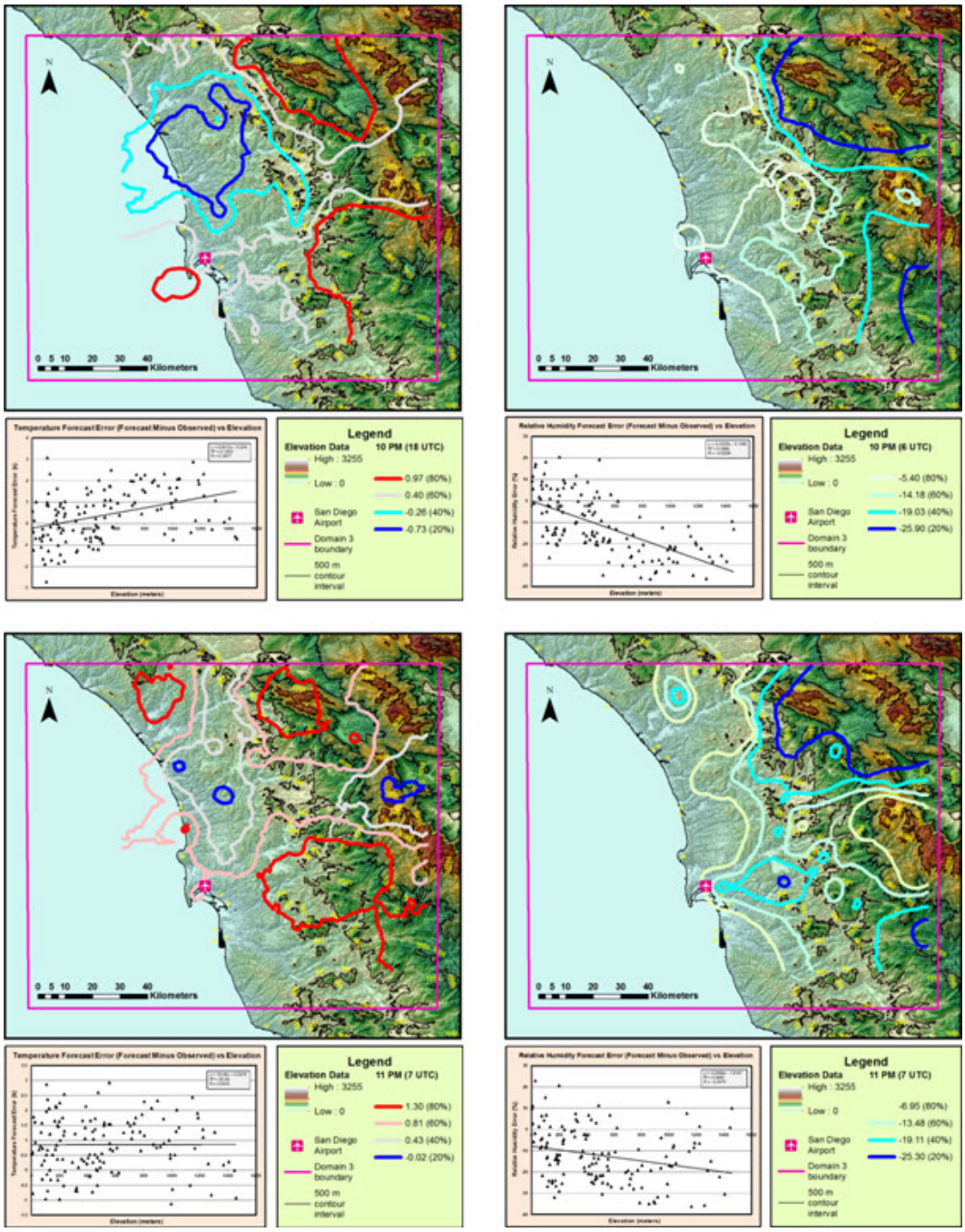

Fig. C-10 February 7, 2012, WRF error (forecast minus observed). Elevation versus temperature forecast error (left column) and elevation versus relative humidity forecast error (right column). The top row is 10:00 PM PST (6 UTC), and the bottom row is 11:00 PM PST (7 UTC).

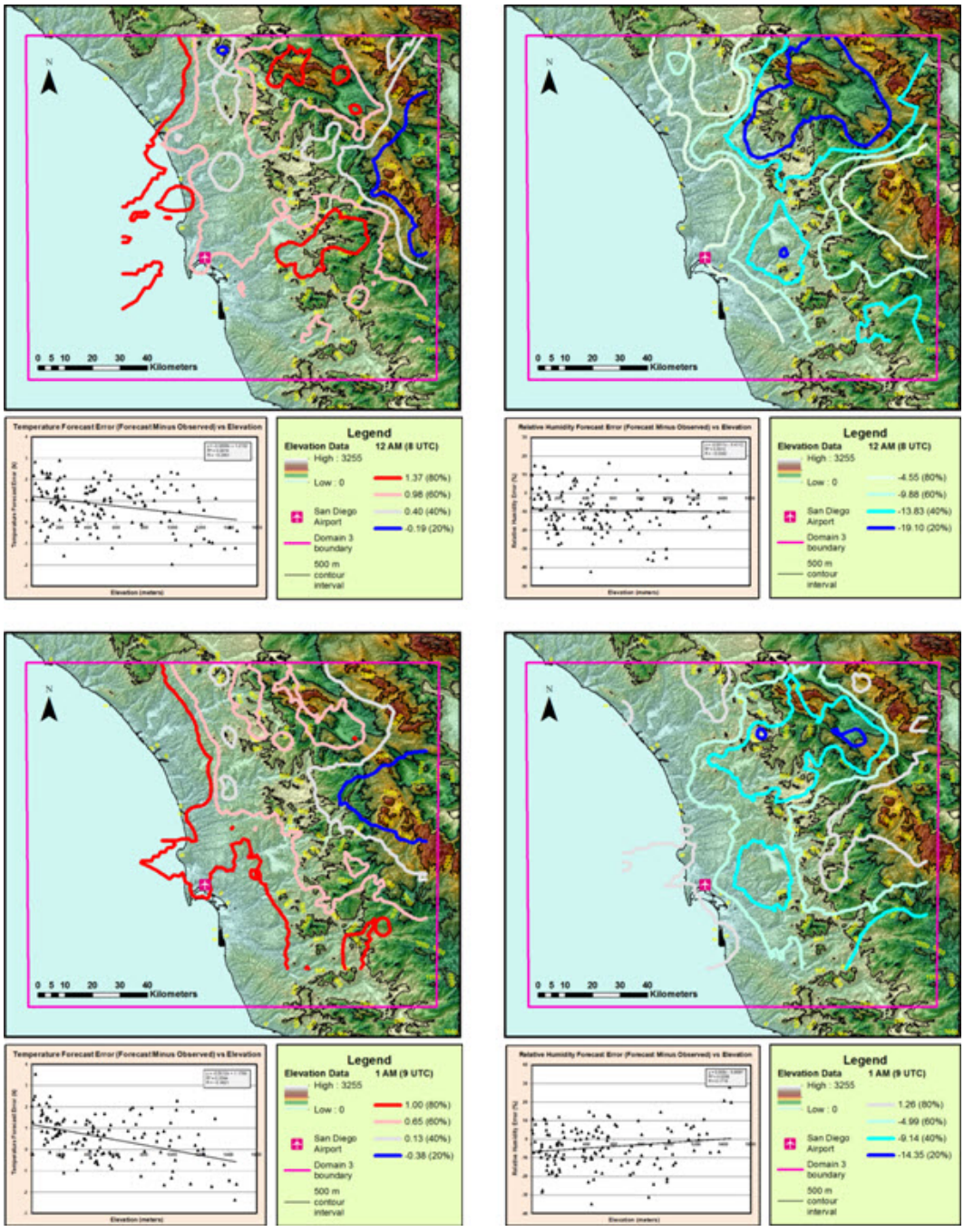

Fig. C-11 February 7, 2012, WRF error (forecast minus observed). Elevation versus temperature forecast error (left column) and elevation versus relative humidity forecast error (right column). The top row is 12:00 AM PST (8 UTC), and the bottom row is 1:00 AM PST (9 UTC).

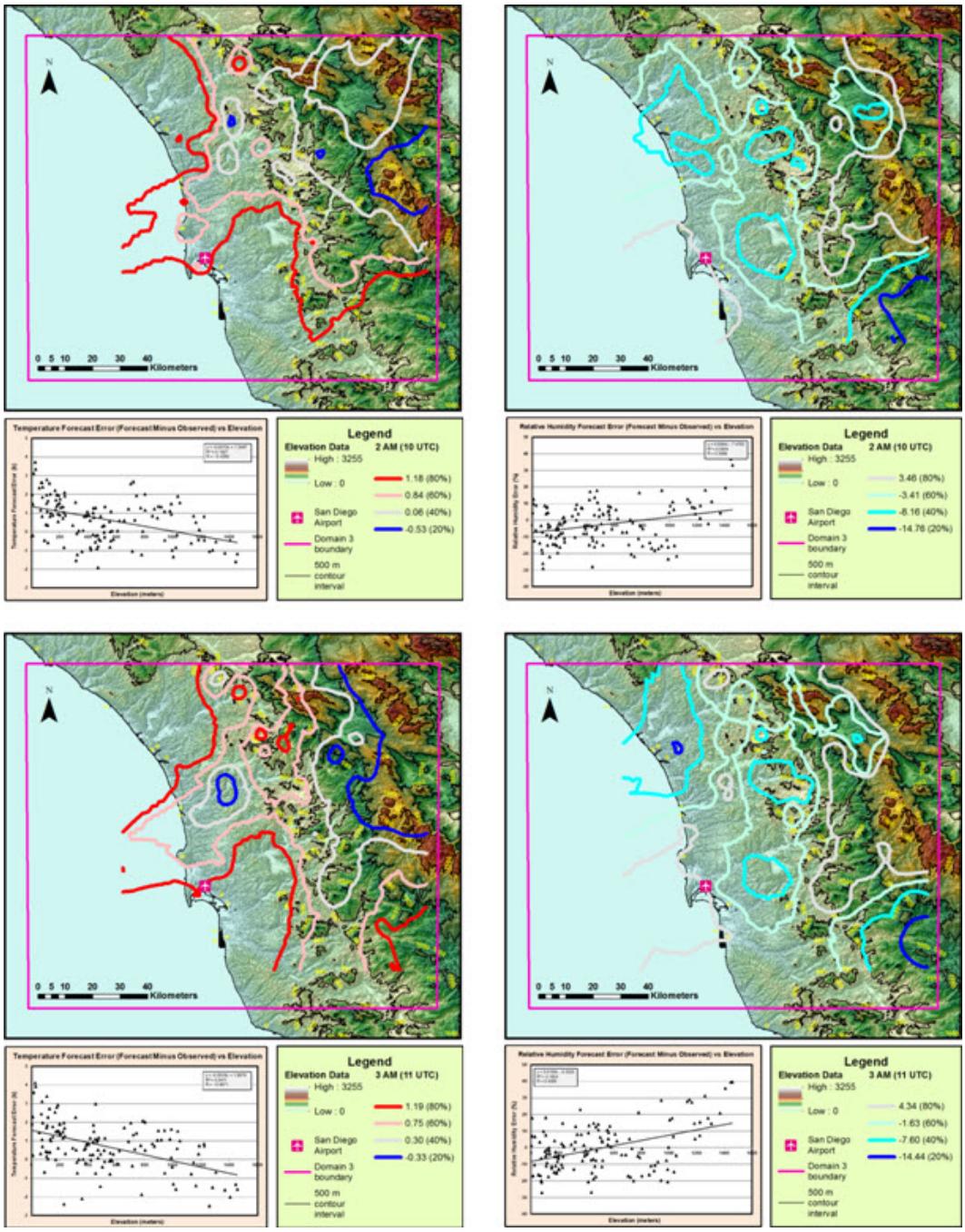

Fig. C-12 February 7, 2012, WRF error (forecast minus observed). Elevation versus temperature forecast error (left column) and elevation versus relative humidity forecast error (right column). The top row is 2:00 AM PST (10 UTC), and the bottom row is 3:00 AM PST (11 UTC).

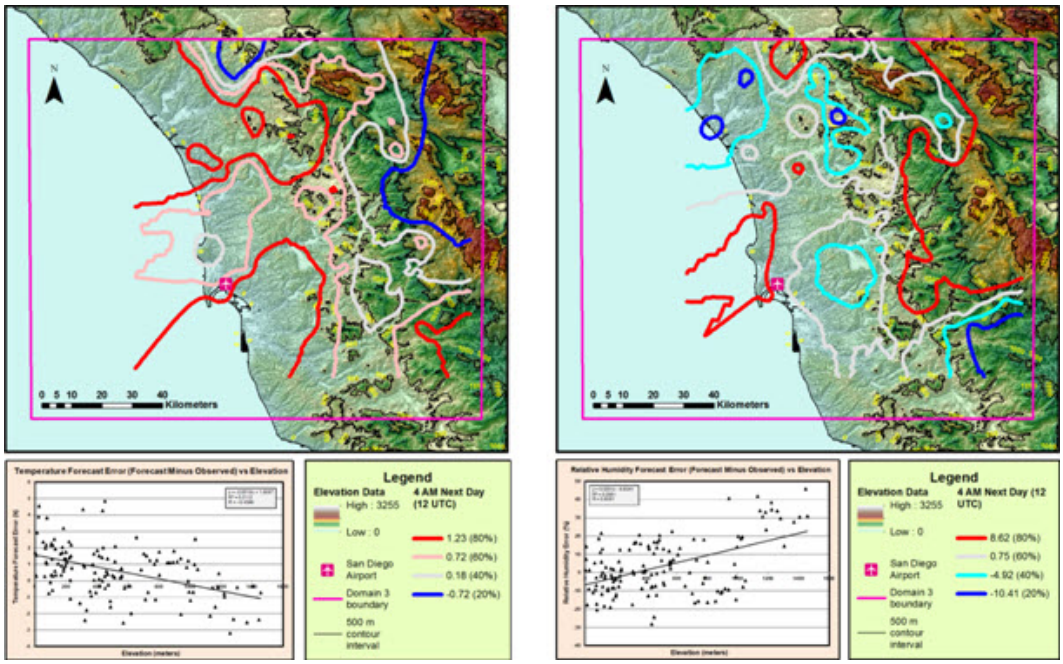

Fig. C-13 February 7, 2012, WRF error (forecast minus observed). 4:00 AM (12 UTC) the next day, elevation versus temperature forecast error (left) and elevation versus relative humidity forecast error (right).

List of Symbols, Abbreviations, and Acronyms

3-D	3-Dimensional
ACARS	Aircraft Communications Addressing and Reporting System
AGL	above ground level
ARL	US Army Research Laboratory
DEM	digital elevation model
EBK	Empirical Bayesian Kriging
FDDA	Four Dimensional Data Assimilation
GFS	Global Forecast System
GIS	Geographic Information System
h	hour(s)
hPa	hectopascal
K	Degrees Kelvin
km	kilometer
LTC	Lieutenant Colonel
MADIS	Meteorological Assimilation Data Ingest System
MET	Model Evaluation Tools
MODE	Method for Object-based Diagnostic Evaluation
MYJ	Mellor–Yamada–Janjić
MyWIDA	My Weather Impacts Decision Aid
NCAR	National Center for Atmospheric Research
NOAA	National Oceanic and Atmospheric Administration
NWP	Numerical Weather Prediction
PBL	Planetary Boundary Layer
PST	Pacific Standard Time
Q-Q	quantile-quantile

RRTM	Rapid Radiative Transfer Model
TAMDAR	Tropospheric Airborne Meteorological Data Reporting
USGS	US Geological Survey
UTC	coordinated universal time
WRE–N	Weather Running Estimate–Nowcast
WRF	Weather Research and Forecasting
WRF–ARW	Weather Research and Forecasting–Advanced Research WRF
WSMR	White Sands Missile Range

1 (PDF)	DEFENSE TECHNICAL INFORMATION CTR DTIC OCA
2 (PDF)	DIRECTOR US ARMY RSRCH LAB RDRL CIO LL IMAL HRA MAIL & RECORDS MGMT
1 (PDF)	GOVT PRINTG OFC A MALHOTRA
4 (PDF)	DIRECTOR US ARMY RSRCH LAB RDRL CIE M T FOLEY J RABY B REEN J SMITH